

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 14:39:25 ; Search time 3772.55 seconds

(without alignments)
10929.928 Million cell updates/sec

Title: US-09-806-276A-4

Perfect score: 2546

Sequence: 1 cgggctagccggagacc.....cagctanaacagggcagta 2546

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_hc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	882.8	34.7	1032	13	BM458203 AGENCOURT
2	879.2	34.5	930	9	AL545294
3	872.4	34.3	938	14	BM804771 AGENCOURT
4	854.6	33.6	920	9	AL547013
5	850.4	33.4	899	9	AL520266
6	844.6	33.2	1968	11	AK019447 Mus muscu

7	844	33.2	1970	11	AK014281
8	842.2	33.1	887	9	AL572131 Mus muscu
9	818	32.1	852	9	AL545262
10	795.8	31.3	889	13	BI916234
11	789.8	31.0	870	13	BI760499
12	783.6	30.8	827	13	BI918195
13	783	30.8	869	14	BQ680912
14	774	30.4	786	13	BI915037
15	770.2	30.3	959	14	BQ878634
16	767.8	30.2	925	13	BI906659
17	763.6	30.0	1487	11	AK012309
18	761	29.9	918	13	BI758178
19	751.8	29.5	866	13	BI861738
20	748.4	29.4	773	13	BI908681
21	744.6	29.2	1031	14	BQ067953
22	725	28.5	808	13	BI767600
23	724.6	28.5	879	14	BQ224749
24	708	27.8	920	14	BQ961461
25	702	27.6	720	14	BM981107
26	701	27.5	876	14	BQ435221
27	697.8	27.4	705	12	BG703895
28	678.6	26.7	1160	12	BG036153
29	671.4	26.4	952	13	BI912055
30	668.8	26.3	704	13	BI760730
31	662.8	26.0	1123	12	BG386157
32	659.4	25.9	685	14	BM990899
33	655.4	25.7	664	13	BI765476
34	649.2	25.5	750	12	BG750650
35	646.8	25.4	1122	14	BQ055733
36	642.4	25.2	1031	12	BG027198
37	636.8	25.0	703	12	BG722248
38	635.8	25.0	927	14	BQ228776
39	630	24.7	631	10	BE385516
40	626.4	24.6	959	12	BG122190
41	625.4	24.6	744	13	BI912364
42	625	24.5	909	12	BE889390
43	623.8	24.5	811	12	BF305603
44	613.8	24.1	700	9	AL544300
45	610.8	24.0	645	10	BE206781

ALIGNMENTS

RESULT 1
BM458203
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BM458203
AGENCOURT_6411553 NIH_MGC_71 Homo sapiens
5', mRNA sequence.
BM458203
BM458203.1 GI:18507243
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1032)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12210 row: n column: 17
High quality sequence stop: 629.
Location/Qualifiers
1..1032

AK014281 Mus muscu
AL572131 AL572131
AL545262 AL545262
BI916234 603178134
BI760499 603045037
BI918195 603183146
BQ680912 AGENCOURT
BI915037 603177226
BQ878634 AGENCOURT
BI906659 603084343
AK012309 Mus muscu
BI758178 603029815
BI861738 603391135
BI908681 603066311
BQ067953 AGENCOURT
BI767600 603080826
BQ224749 AGENCOURT
BQ961461 AGENCOURT
BM981107 UT-CF-EN1
BQ435221 AGENCOURT
BQ703895 602686951
BG036153 602326182
BI912055 603069052
BG386157 602455339
BM990899 UT-H-D10-
BI765476 603050539
BG750650 602708506
BQ055733 AGENCOURT
BQ027198 602295884
BG722248 602693416
BQ228776 AGENCOURT
BE385516 601278030
BG122190 602349739
BI912364 603067583
BE889390 601513281
BF305603 601893434
AL544300 AL544300
BE206781 ba03011.y


```
Db 339 TTGAGAAAATGATGTGACAGTCTCTATCAAAAGGATTCAGAAACATATACCATCTGTG 398
Qy 361 AAGAAAGTGGCCCTTCTCCGCTTGCNAATAGACATCTCAATTCCTCAAAATGCCAGC 420
Db 399 AAGAAAGTGGCCCTTCTCCGCTTGCNAATAGACATCTCAATTCCTCAAAATGCCAGC 458
Qy 421 CAAGACCCCAATTTACCTGAAAGCAGCAGCAATTAACAAGAGGAAAGAAATTAACACTGAG 480
Db 459 CAAGACCCCAATTTACCTGAAAGCAGCAGCAATTAACAAGAGGAAAGAAATTAACACTGAG 518
Qy 481 GGACATCTGTCTCTGATATGATCAGTCCCGCTTGGAGACTTTGCGCCACACCATCCA 540
Db 519 GGACATCTGTCTCTGATATGATCAGTCCCGCTTGGAGACTTTGCGCCACACCATCCA 578
Qy 541 CATTGGCAAGAGGCGCAGCAGTCTTTGGAGATATTCCTTCTCAAGGAACTA 600
Db 579 CATTGGCAAGAGGCGCAGCAGTCTTTGGAGATATTCCTTCTCAAGGAACTA 638
Qy 601 CGAGCTTTTACCTGAAACACAGGAGAAAGCACACCTGGGCCAGTTCCTGGGCAATAATGA 660
Db 639 CGAGCTTTTACCTGAAACACAGGAGAAAGCACACCTGGGCCAGTTCCTGGGCAATAATGA 698
Qy 661 GTTCTTCGGGGCAACAGACCTCGGACTCTGTGTTACAGAAAGCCCTCCCGGTGCT 720
Db 699 GTTCTTCGGGGCAACAGACCTCGGACTCTGTGTTACAGAAAGCCCTCCCGGTGCT 758
Qy 721 CAAAATGCCATCTCCCTCCGACCATTTGGAGGATCCCAAGCTCTCATGTTGCCCTTAT 780
Db 759 CAAAATGCCATCTCCCTCCGACCATTTGGAGGATCCCAAGCTCTCATGTTGCCCTTAT 818
Qy 781 GTCCACGAGTACATTTAATTCACACAGGAGTCTCTCGGGCCAGCAAGCTGCCAGGCT 840
Db 819 GTCCACGAGTACATTTAATTCACACAGGAGTCTCTCGGGCCAGCAAGCTGCCAGGCT 878
Qy 841 TAGCTGGAGCCGCTCATGGAGGAAAAAGCTCAGAGAGAAAGCAGTCTGTTGG 893
Db 879 TAGCTGGAGCCGCTCATGGAGGAAAAA-CTCAGGAGAAAAAFCAGTCTGTTGG 930

RESULT 3
BM804771
LOCUS
DEFINITION
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12716 row: f column: 18
High quality sequence stop: 659.
Location/Qualifiers
1. .938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5725457"
/clone_lib="NIH_MGC_125"
```

```
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site:1: EcoRV (destroyed); Site:2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."
BASE COUNT 230 a 269 c 249 g 189 t 1 others
ORIGIN
Query Match 34.3%; Score 872.4; DB 14; Length 938;
Best Local Similarity 98.5%; Pred. No. 2e-167;
Matches 901; Conservative 0; Mismatches 12; Indels 2; Gaps 2;
Qy 1 CGGGGCTAGCCCGGAGACCCGCGGCTCGGGGCGCTTCACGCCGCTCTCGGAGCGG 60
Db 25 CGGGGCTAGCCCGGAGACCCGCGGCTCGGGGCGCTTCACGCCGCTCTCGGAGCGG 84
Qy 61 ATATGCGGTGAGCAGGACCCAGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 120
Db 85 ATATGCGGTGAGCAGGACCCAGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 144
Qy 121 GATTGCGTGCCTCGGCGGAGGCGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 180
Db 145 GATTGCGTGCCTCGGCGGAGGCGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 204
Qy 181 TGTCCTGGTGTGCTCAAGACTGGATGATGATGATGATGATGATGATGATGATGATG 240
Db 205 TGTCCTGGTGTGCTCAAGACTGGATGATGATGATGATGATGATGATGATGATGATG 264
Qy 241 CGTAACCTCAGGAGGTTCTTTGACCCCATCTCATTTGCAAGCCACTTCTCAAGCCAC 300
Db 265 CGTAACCTCAGGAGGTTCTTTGACCCCATCTCATTTGCAAGCCACTTCTCAAGCCAC 324
Qy 301 TTGAGAAAAATGATGTGACAGTTCCTATCAAAAAGGATTCAGAAACATATACCATCTGTG 360
Db 325 TTGAGAAAAATGATGTGACAGTTCCTATCAAAAAGGATTCAGAAACATATACCATCTGTG 384
Qy 361 AAGAAAGTGGCCCTTCTCCGCTTGCNAATAGACATCTCAATTCCTCAAAATGCCAGC 420
Db 385 AAGAAAGTGGCCCTTCTCCGCTTGCNAATAGACATCTCAATTCCTCAAAATGCCAGC 444
Qy 421 CAAGACCCCAATTTACCTGAAAGCAGCAGCAATTAACAAGAGGAAAGAAATTAACACTGAG 480
Db 445 CAAGACCCCAATTTACCTGAAAGCAGCAGCAATTAACAAGAGGAAAGAAATTAACACTGAG 504
Qy 481 GGACATCTGTCTCTGATATGATCAGTCCCGCTTGGAGACTTTGCGCCACACCATCCA 540
Db 505 GGACATCTGTCTCTGATATGATCAGTCCCGCTTGGAGACTTTGCGCCACACCATCCA 564
Qy 541 CATTGGCAAGAGGCGCAGCAGTCTTTGGAGATATTCCTTCTCAAGGAACTA 600
Db 565 CATTGGCAAGAGGCGCAGCAGTCTTTGGAGATATTCCTTCTCAAGGAACTA 624
Qy 601 CGAGCTTTTACCTGAAACACAGGAGAAAGCACACCTGGGCCAGTTCCTCGGCAATAATGA 660
Db 625 CGAGCTTTTACCTGAAACACAGGAGAAAGCACACCTGGGCCAGTTCCTCGGCAATAATGA 684
Qy 661 GTTCTTCGGGGCAACAGCAGTCTGTGTTTTCAGAGAAAGCCCTCCCGGTGCT 720
Db 685 GTTCTTCGGGGCAACAGCAGTCTGTGTTTTCAGAGAAAGCCCTCCCGGTGCT 744
Qy 721 CAAAATGCCATCTCCCTCCGACCATTTGGAGGATCCCAAGCTCTCATGTTGCCCTTAT 780
Db 745 CAAAATGCCATCTCCCTCCGACCATTTGGAGGATCCCAAGCTCTCATGTTGCCCTTAT 804
Qy 781 GTCCACGAGTACATTTAATTCACACAGGAGTCTCTCGGGCCAGCAAGCTGCCAGGCT 840
Db 805 GTCCACGAGTACATTTAA-TCCCAACAGGAGTCTCTCGGGCCAGCAAGCTGCCAGGCT 863
```


GAGAGAGATTCTCGAGTAATAATAATTAATCCCGCCCCCCCC 3'. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES

source

Location/Qualifiers

1. 1970
/organism="Mus musculus"
/strain="C57BL/6j"
/db_xref="FANTOM_DB:3200001F04"
/db_xref="MGD:1898924"
/db_xref="taxon:10090"
/clone="3200001F04"
/tissue="type:head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="14, 17 days embryo"
505..1269
/note="data source:SPTR, source key:Q9UK12, evidence:ISS
homolog to CRIB-CONTAINING BORG2 PROTEIN
putative"

CDS

BASE COUNT 517 a 490 c 467 g 496 t
ORIGIN
Query Match 33.2%; Score 844; DB 11; Length 1970;
Best Local Similarity 70.8%; Pred. No. 8.8e-162;
Matches 1325; Conservative 0; Mismatches 465; Indels 82; Gaps 12;
QY 20 CGGCCACCGGCTGGGGCCCTTCACGCGCTCTCGGAGCGGATATGCGGTGAGCAGCA 79
DB 109 CGGCTGCTCCCGGAGAGCGCGCGAGTCCCGGAGCAGATTCGCGGTGAACCGTTG 168
QY 80 CCACGCGGACACTCGGCTGATCTCGG-----CACAGCGGACGAGATTGCGT 128
DB 169 CGCGCCCTCTGTTCTGCGCGGACCGAGGAGGAGTGTTCGCGCCCTCGGAGCGACCCGG 228
QY 129 GCGCCGCGGAGGCGCGGCGGAGCGGCTGCGATCTCAGCGGCGGCGGTTCTGCTCGT 188
DB 229 GCGCGCGGCGGCGGCTCCGCGGAGAGCTAGGAGCTTAGAGCTGCTGTTCTCTGAG 288
QY 189 TTGCTGCTCAAGACTGGATGATGTAAGTAACTGGCTCTAGGAGCGCTCACTTGGCCGTAACCT 248
DB 289 CTGTGGAAGGAGCTCGATAAGTAACTGGCTCTCGGAGAGACCGCGGTGGCCCTGCGCT 348
QY 249 CAGGAAGTTCCTTTTGACCCCATCTCATTTTGAAGCCACTTCGAAGCCACTTGAGAAA 308
DB 349 GAGCAAGATCTCCAAGACCTCTCTCGTTTTCAGAGCCACTTGAGAAG-----A 397
QY 309 AATGATGTGACAGTTCCTATCAAAAAGCA---TTCAAGAAATATACCATCTGTGAAGAA 365
DB 398 AATGATGTGACAAATTCCTATGAAAAGGAGCGCTTTTGAAGTGTAGGATACCGCGTGAGAA 457
QY 366 AGTGGCCCTTTCTCCCGCTTCGAAATAGACATCTCAAAATTCGAAATTCGAGCCGCAAGA 425
DB 458 GGTCACTCTTCCCTGGTCTGTGAATAGACATTCCTCAAAATTCGAAATTCGCTGCGCAAGA 517
QY 426 CCCCATTTCCTGAAAGCGCCATTAACAAGAAAGGAGAAATTTAAACTGAGGAGCA 485
DB 518 CCCCATTTCCTGAAAGCGCTTAACAAGAAAGGAGAAATTTAAACTGAGGAGCA 577
QY 486 TTCTGCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 545
DB 578 TTCTGCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 637
QY 546 GCAAGAGGCGGACGAGTGTCTTTTGAGATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 605
DB 638 GCAAGAGGCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 697

QY 606 TTTTACCTGGAAACAGGAGAAAGCACACCTGGCCAGTTCCTGGCAGTAATGACTTCT 665
DB 698 TCTTGGCGGAAACCAAGAGAGCCACACTCTGCCAGTTCCTGGGCGCAATGACTTCT 757
QY 666 TCCGGGCAACAGCAGCCTCGGACTCTGTGTTACAGAAAGCGCCTCCCGGCTCTCAAAA 725
DB 758 TCCGGGCAACAGCAGCCTCGGACTCAATGTTTACAGAAACACCTCCCGGCTCTCAAAA 817
QY 726 ATGCCATCTCCCTCCCGACCATTTGGAGGATCCCAAGCTCTCATGTTGCGCTTATGTCAC 785
DB 818 ATGCCATCTCCCTCCCGACCATTTGGGAGATCCAGGCTCTGATGCTGCTTATTTGCC 877
QY 786 CAGTGACATTTAATTCACAGAGTCTCTTCGGGCGCAGCAAGAGTGCAGGCTTAGCT 845
DB 878 CAGTGACATTTAATTCACAGAGTCTCTTTGGGCGTCTTAAGTTGCCAGGCTCAGCT 937
QY 846 GCGAGCCGTCATGGAGGAAAGAGCTCAGGAGAAAGCAGTCTGTTGGAGAAATGGACAG 905
DB 938 GTGAGCTGTCTGGAAGAGAAAGTTTCAGGAGCAAGTAGTAGTATTAGAGAAATGGGCGAG 997
QY 906 TCCACAGGAGAGACCTCGTGGGCTCCAGCGGCTTCTGCATCTCAGTCCAGCCCAAGCA 965
DB 998 TCCACAGGAGAGACCTCGTGGGCTCCAGTGGTTCGGGCTCTCAGTCCAGCCAGGCGCC 1057
QY 966 GAGACAGCCACTCTCCAGCCTGTCGGAACAGTACCCGAGTGGCCAGCCAGGACATGT 1025
DB 1058 GCGACAGCCACTCTCCAGCCTGTCGGAACAGTCTCCGAGTGGCCGCGCAGCAGCATGT 1117
QY 1026 TTGACCATCCACCCCATCGAGCTCATCAAGGGAAGAGTAAAGTTCAGAGAGTCCCTCT 1085
DB 1118 TTGAGCATCTGCTGCTGAGCTGTGAAGTCAAGAGCTAAATTCAGAGAGTCCCTCT 1177
QY 1086 CTGACCTTACAGGTTCCCTCTCTCCCTGAGCTTTGTTGGGCGCTCACTTTTGGATG 1145
DB 1178 CTGACCTGACGCGCTCCCTGCTCTCCCTGAGTCTGATCTTGGGCTTCACTTTTGGATG 1237
QY 1146 AGTGTCTGAATTAATGATATAAATAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1205
DB 1238 AGTGTCTGAATTAATGATATAAATAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1297
QY 1206 AAAAGTAG-----AAAAACAACCTAACACAGTTCGAAGAGAGG 1246
DB 1298 AAGAGGTACAAAATAAGCAGAAAAACAATAAATAAATAAATAAATAAATAAATAAATAA 1357
QY 1247 CTTCGGGAGCTG---TATTTGAGTGTGTTGGGTTTCTTAAATAA-----TATT 1296
DB 1358 CTGCTCTGACTGTTTATTTGAGCGCATGTGATGATTTTCTTAAATAAATAAATAAATAA 1417
QY 1297 CTTACAAAGTATTTTACCTGTTATGCGCTGTTTGCAGAA-----AACAAAT 1344
DB 1418 CTTACAAAGTATTTTACCTGTTATGCGCTGTTTGCAGAA-----AACAAAT 1477
QY 1345 TAGAAAAACAACAAGCAAAACCTATCTTGGCAAAAAAGAGAGTGTGAGTTCAGAGCCCA 1404
DB 1478 TTTAAAAATTTAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1537
QY 1405 TTTTCAGGAGGCAATGTGTGATGTTGCGGCTCACATATTTGTTG--CAGACACACAGAAAT 1462
DB 1538 AAAGAAAAAGAAAAAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1597
QY 1463 CTGCTTGGCGAGATTTGGGCACTAGCTATGTAAGGCGTGTGAGGAGTGTGAGTGTGAGT 1522
DB 1598 CATACATGCTCAGTTTGTGTTGTGAGATGCAGAAATCCAGGCTGTCTAAGAGGCTGGG 1657
QY 1523 CACGGAATTTATAGCACTCCGACATTTCTGAGCAAGAGGAGTGTGAGTGTGAGTGTGAGT 1581
DB 1658 TCCATGGGTCTCTCAAGGCAATTCATGATGTTATGAGAAAGAACTTACTTTATTTGG 1717
QY 1582 CACCTAAGCCCTTTTGTAGACTCTTTCT-----ATATATTGCTTAGGCTCACCATAGC 1635
DB 1718 CACCTAAGCCCTTTTGTAGACTCTTTCTTTTGTATATGTTATCTTTTGGGTTCCACCAC 1777

/clone="CS0DI028YL18"
 /clone_lib="LRI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site.1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 291 a 155 c 142 g 255 t 9 others
 ORIGIN

Query Match 32.1%; Score 818; DB 9; Length 852;
 Best Local Similarity 98.2%; Pred. No. 2.4e-156;
 Matches 820; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

1278 GGTTCCTAAATATATTTACAAAGTATTTTACCTGTTATGCTGTTTGCAAA 1337
 |||||||
 852 GGTTCCTAAATATATTTACAAAGTATTTTACCTGTTATGCTGTTTGCAAA 793
 |||||||

1338 AACATTTAGAAAAAACAAGCAAAACCTATCTTGGCAAAAAAGGAAGTACGA 1397
 |||||||
 792 AACATTTAGAAAAAACAAGCAAAACCTATCTTGGCAAAAAAGGAAGTACGA 733
 |||||||

1398 GAGCCCATTTTCAGGAGCATTGGTGTGCTGCCTACATATTTTGCAGACACAAA 1457
 |||||||
 732 GAGCCCATTTTCAGGAGCATTGGTGTGCTGCCTACATATTTTGCAGACACAAA 673
 |||||||

1458 GAAATCTGGCTGGCCAGGATTGGCAGTATGATGAAGGGCTGAGCGAGTCACATTAAGG 1517
 |||||||
 672 GAAATCTGGCTGGCCAGGATTGGCAGTATGATGAAGGGCTGAGCGAGTCACATTAAGG 613
 |||||||

1518 AACTTCAGGAACTTTATAGCACTCCGCAATTTCTGAGCAAGAGGAAGTCAAAATTTAT 1577
 |||||||
 612 AACTTCAGGAACTTTATAGCACTCCGCAATTTCTGAGCAAGAGGAAGTCAAAATTTAT 553
 |||||||

1578 TTAACACCTAAGCCTTTTGTAGACTCTTTCTATATATTTCTTAGCTCACCATACGGA 1637
 |||||||
 552 TTAACACCTAAGCCTTTTGTAGACTCTTTCTATATATTTCTTAGCTCACCATACGGA 493
 |||||||

1638 ATTCTCAGTGTAAACCTTTCTGTTTTCACATTTGAACCTTTATGGTTTGGGATTT 1697
 |||||||
 492 ATTCTCAGTGTAAACCTTTCTGTTTTCACATTTGAACCTTTATGGTTTGGGATTT 433
 |||||||

1698 TCTCTGAGTCTTATATATCCCTATATATATATATATATATATATATATATATATATAT 1757
 |||||||
 432 TCTCTGAGTCTTATATATCCCTATATATATATATATATATATATATATATATATATAT 373
 |||||||

1758 CTACATGTTGTAAGACACAGCAAGATTTACTGTAACTAAGTTATTTTAAAGTTAAA 1817
 |||||||
 372 CTACATGTTGTAAGACACAGCAAGATTTACTGTAACTAAGTTATTTTAAAGTTAAA 313
 |||||||

1818 ATATATTTTACGTCCTTTGGCTTTTATTTGAGAGCTTACATTTTATAGATTTACAT 1877
 |||||||
 312 ATATATTTTACGTCCTTTGGCTTTTATTTGAGAGCTTACATTTTATAGATTTACAT 253
 |||||||

1878 CAGATGTTGTCATTTATTTCCATTTGGGATTTCCATTTGAAGCTGTATGTCGGTGTGG 1937
 |||||||
 252 CAGATGTTGTCATTTATTTCCATTTGGGATTTCCATTTGAAGCTGTATGTCGGTGTGG 193
 |||||||

1938 AAAAGTGTATTCATCTAGTTTCTTTCTTCATCTGTTATATATATCTTTTAAAGCAAC 1997
 |||||||
 192 AAAAGTGTATTCATCTAGTTTCTTTCTTCATCTGTTATATATCTTTTAAAGCAAC 133
 |||||||

1998 CAATTAACGGATTTAAAGTGTAAAGGCACAGGTTTACTCATGATGCTTCTGCAGAGCTGT 2057
 |||||||
 132 CAATTAACGGATTTAAAGTGTAAAGGCACAGGTTTACTCATGATGCTTCTGCAGAGCTGT 73
 |||||||

2058 GGGCTACACCATATGTTATTTTGGAAATATAGTATTTTAGTACAGTACATACT 2112
 |||||||

Db 72 GGGCTACACATATGTTATTTGGAAATATAGTATTTTAGTACAGTACATACT 18
 |||||||

RESULT 10
 BI916234 889 bp mRNA linear EST 16-OCT-2001
 LOCUS 603178134F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5242305 5',
 DEFINITION mRNA sequence.
 ACCESSION BI916234
 VERSION BI916234.1 GI:16180195
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 889)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1611 row: c column: 10
 High quality sequence stop: 865.
 Location/Qualifiers
 1. .889
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5242305"
 /clone_lib="NIH_MGC_121"
 /lab_host="DH10B"
 /notes="Organ: brain; Vector: pCMV-SPORT6; Site.1: NotI;
 Site.2: EcoRV (destroyed); RNA source anonymous pool of 3
 fetal brains, female age 20 weeks, female age 24 weeks,
 and male age 26 weeks. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 0.7-3.5 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 017. Note:
 this is a NIH_MGC Library."

BASE COUNT 214 a 264 c 226 g 185 t
 ORIGIN

Query Match 31.3%; Score 795.8; DB 13; Length 889;
 Best Local Similarity 98.5%; Pred. No. 7.8e-152;
 Matches 846; Conservative 0; Mismatches 7; Indels 6; Gaps 4;

QY 1 CGGGGCTAGCCCGGAGACCCGCCACCGCCCTGGGCGCTTCACGCCGCTCTCGGAGCGG 60
 |||||||

Db 29 CGGGGCTAGCCCGGAGACCCGCCACCGCCCTGGGCGCTTCACGCCGCTCTCGGAGCGG 88
 |||||||

QY 61 ATAATCGGTGAGCAGCAGCACCGCGGCGAGACTCGGCTGGGATCTGCGCACAGCGGAGG 120
 |||||||

Db 89 ATAATCGGTGAGCAGCAGCACCGCGGCGAGACTCGGCTGGGATCTGCGCACAGCGGAGG 148
 |||||||

QY 121 GATTGCGTGGCCCGGAGGAGCCCGGGGAGCGGCTGGGATCTCTAGGAAGCCTCAGTGGC 180
 |||||||

Db 149 GATTGCGTGGCCCGGAGGAGCCCGGGGAGCGGCTGGGATCTCTAGGAAGCCTCAGTGGC 208
 |||||||

QY 181 TGTCTCTGTTGTGGTCAAGTGGATGATGTAACCTGGCTCTCTAGGAAGCCTCAGTGGC 240
 |||||||

Db 209 TGTCTCTGTTGTGGTCAAGTGGATGATGTAACCTGGCTCTCTAGGAAGCCTCAGTGGC 268
 |||||||

QY 241 CGTAACTCAGGAAGGTTCTCTTTTGACCCCATCTATTTTCGAAGCCACTTCTGAAGCCAC 300
 |||||||

```

Db      269  CGTAACCTCAGGAAGGTTCTCTTTGACCCCATCTCATTTGGAAGCCACTTCTGAAGCCAC 328
QY      301  TTGAGAAAATGATGTGACAGTCTCTATCAAAAAGGATTCAGAACATATACCATCTGTG 360
Db      329  TTGAGAAAATGATGTGACAGTCTCTATCAAAAAGGATTCAGAACATATACCATCTGTG 388
QY      361  AGAAGAGTGCCCTTTCTCCCGCTTGCAAAATAGACATCTCAAAATCCAAAATGCCAGC 420
Db      389  AGAAGAGTGCCCTTTCTCCCGCTTGCAAAATAGACATCTCAAAATCCAAAATGCCAGC 448
QY      421  CAAGACCCCAATTTACCTCAAGACGACCAATAACAAGAAAGAAAGAAATTTAAACTGAG 480
Db      449  CAAGACCCCAATTTACCTCAAGACGACCAATAACAAGAAAGAAAGAAATTTAAACTGAG 508
QY      481  GGACATCTCTCTCTGATATGATCAGTCCCGCTTGGAAGTTCGCGACATTCGCGACACCATCCA 540
Db      509  GGACATCTCTCTCTGATATGATCAGTCCCGCTTGGAAGTTCGCGACATTCGCGACACCATCCA 568
QY      541  CATTGGCAAGAGGCCGACGAGTCTCTTTGGAGATATTTCTTCTTCAAGGGAACATA 600
Db      569  CATTGGCAAGAGGCCGACGAGTCTCTTTGGAGATATTTCTTCTTCAAGGGAACATA 628
QY      601  CGAGCTTTTACCTGGAACACAGGAGAACACACCTGGCCAGTTCCTCTGGGCATATGA 660
Db      629  CGAGCTTTTACCTGGAACACAGGAGAACACACCTGGCCAGTTCCTCTGGGCATATGA 688
QY      661  GTTCTTCGGGGCCAAACAGCCTCGGACTCTGTGTTTCAAGAAAGCCCTCCCGGTGCT 720
Db      689  GTTCTTCGGGGCCAAACAGCCTCGGACTCTGTGTTTCAAGAAAGCCCTCCCGGTGCT 748
QY      721  CAAAATGCCATCT-CCCTCCGACATTTGAGAGATCCCAAGTCTCATGTTGCCCTTAT 779
Db      749  CAAAATGCCATCTCCCTCCGACATTTGAGAGATCCCAAGTCTCATGTTGCCCTTAT 808
QY      780  TGTCAACAGTGACATTTAA--TTCCAAACAGGAGTCTTC--GGGCCAGCAAGTGGCC 835
Db      809  TGTCAACAGTGACATTTAAATTCAAAACAGGAGTCTTC--GGGCCAGCAAGTGGCC 868
QY      836  AGG-CTTAGCTGGAGGCC 853
Db      869  AGGCTTAGCTGGAGGCC 887

```

RESULT 11

```

LOCUS      BI760499
DEFINITION 603045037f1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5185256 5',
            mRNA sequence.
VERSION    BI760499
KEYWORDS   EST.
SOURCE     human.

```

ORGANISM

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1462 row: j column: 09
High quality sequence stop: 865.
Location/Qualifiers
I. .870
/organism="Homo sapiens"

```

FEATURES

```

source

```

```

/db_xref="taxon:9606"
/clone="IMAGE:5185256"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/Note="Organ: pooled colon, kidney, stomach; Vector:
pcmv-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."
BASE COUNT 212 a 247 c 228 g 182 t 1 others
ORIGIN

```

Query Match

```

Best Local Similarity 31.0%; Score 789.8; DB 13; Length 870;
Matches 819; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

```

```

QY      1  CGGGCTAGCCCGGAGACCGCCACCGGCTGGGGCGCTTCAGCGCGTCTGGAGCGG 60
Db      24  CGGGCTAGCCCGGAGACCGCCACCGGCTGGGGCGCTTCAGCGCGTCTGGAGCGG 83
QY      61  ATATGCGGTGAGCAGGACACCGCGGACACTCGGCTGGATCTGCGACACGCGCAGG 120
Db      84  ATATGCGGTGAGCAGGACACCGCGGACACTCGGCTGGATCTGCGACACGCGCAGG 143
QY      121  GATTGCGTGGCGGAGGAGCGGCGGCGGCTGGATCTTCAGCGGCGCGGTT 180
Db      144  GATTGCGTGGCGGAGGAGCGGCGGCGGCTGGATCTTCAGCGGCGCGGTT 203
QY      181  TGTCTGTTGTGCTCAAGACTGGATGATGTAATCTCTCTCTCTCTCTCTCTCTCT 240
Db      204  TGTCTGTTGTGCTCAAGACTGGATGATGTAATCTCTCTCTCTCTCTCTCTCTCT 263
QY      241  CGTAACCTCAGGAAGTCTCTTTGACCCCATCTCATTTGGAAGCCACTTCTGAAGCCAC 300
Db      264  CGTAACCTCAGGAAGTCTCTTTGACCCCATCTCATTTGGAAGCCACTTCTGAAGCCAC 323
QY      301  TTGAGAAAATGATGTGACAGTTCCTATCAAAAAGGATTCAGAAAATATACCATCTGTG 360
Db      324  TTGAGAAAATGATGTGACAGTTCCTATCAAAAAGGATTCAGAAAATATACCATCTGTG 383
QY      361  AAGAAAGTGCGCTTTCTCCCGCTTGCAAAATAGACATCTCAAAATCCAAAATGCCAGC 420
Db      384  AAGAAAGTGCGCTTTCTCCCGCTTGCAAAATAGACATCTCAAAATCCAAAATGCCAGC 443
QY      421  CAAGACCCCAATTTACCTGAAAGCAGCAATACAGAAAGAAAGAAATTTAAACTGAG 480
Db      444  CAAGACCCCAATTTACCTGAAAGCAGCAATACAGAAAGAAAGAAATTTAAACTGAG 503
QY      481  GGACATCTCTCTCTGATATGATCAGTCCCGCTTGAGACTTTTGGACACCATCCA 540
Db      504  GGACATCTCTCTCTGATATGATCAGTCCCGCTTGAGACTTTTGGACACCATCCA 563
QY      541  CATTGGCAAGAGGCCGACGAGTCTCTTTGGAGATATTTCTTCTTCAAGGGAACATA 600
Db      564  CATTGGCAAGAGGCCGACGAGTCTCTTTGGAGATATTTCTTCTTCAAGGGAACATA 623
QY      601  CGAGCTTTTACCTGGAACACAGGAGAACACACCTGGGCGAGTTCCTCTGGGCATATGA 660
Db      624  CGAGCTTTTACCTGGAACACAGGAGAACACACCTGGGCGAGTTCCTCTGGGCATATGA 683
QY      661  GTTCTTCGGGGCCAAACAGCCTCGGACTCTGTGTTTCAAGAAAGCCCTCCCGGTGCT 720
Db      684  GTTCTTCGGGGCCAAACAGCCTCGGACTCTGTGTTTCAAGAAAGCCCTCCCGGTGCT 743
QY      721  CAAAATGCCATCTCCCTCCCGACATTTGGAGGATCCCAAGCTCTCATGTTGCCCTTAT 780
Db      744  CAAAATGCCATCTCCCTCCCGACATTTGGAGGAT-CCAAGCTCTCATGTTGCCCTTA-T 801

```


QY 781 GTCACCACTGACATTTAATCCAAACAGGAGTCTTCGGGCCACCAAGCTGCC 834
 Db 802 GTCACCACTGACATTTAATCCAAACAGGAGTCTTCGGGCCACCAAGCTGCC 855

RESULT 12

BI918195 827 bp mRNA linear EST 16-OCT-2001
 LOCUS 603183146F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5247130 5',
 DEFINITION mRNA sequence.

ACCESSION BI918195
 VERSION BI918195.1 GI:16181952
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 827)

NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11623 row: 1 column: 11

High quality sequence stop: 766.

FEATURES

source

1..827

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5247130"

/clone_lib="NIH_MGC_121"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: EcoRV (destroyed); RNA source anonymous pool of 3

fetal brains, female age 20 weeks, female age 24 weeks,

and male age 26 weeks. Library is oligo-dT primed and

directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.7 kb, insert size range

0.7-3.5 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 017. Note:

this is a NIH-MGC Library."

204 a 237 c 213 g 173 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 30.8%; Score 783.6; DB 13; Length 827;

Matches 819; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY 1 CGGGGCTAGCCGAGACCCGCGCCACCGCGCTGGGGCGCTTCAGCGCGTCTCGAGCGG 60

Db 4 CGGGGCTAGCCGAGACCCGCGCCACCGCGCTGGGGCGCTTCAGCGCGTCTCGAGCGG 63

QY 61 ATATGCGGTGACGAGCACCACCGCGGAGACTCGGCTGGGATCTGCGCACAGCGGAGG 120

Db 64 ATATGCGGTGACGAGCACCACCGCGGAGACTCGGCTGGGATCTGCGCACAGCGGAGG 123

QY 121 GATTGCGTGGCGCGCGGAGCGCGGCGGAGCGGCTGGGATCTCTAGCGGCGCGGTT 180

Db 124 GATTGCGTGGCGCGCGGAGCGCGGCGGAGCGGCTGGGATCTCTAGCGGCGCGGTT 183

QY 181 TGCTCTGGTGTGCTGAAGACTGGATGATGATGATGATGATGATGATGATGATGATGAT 240

Db 184 TGCTCTGGTGTGCTGAAGACTGGATGATGATGATGATGATGATGATGATGATGATGAT 243

QY 241 CGTAACCTCAGGAAGTTCCTTTTGACCCCATCTCAATTCGAAGCCACTTCTGAAGCCAC 300
 Db 244 CGTAACCTCAGGAAGTTCCTTTTGACCCCATCTCAATTCGAAGCCACTTCTGAAGCCAC 303
 QY 301 TTGAGAAAATGATGTGACAGTTCCTATCAAAAAGGATTTCAGAAAACATATACCATCTGTG 360
 Db 304 TTGAGAAAATGATGTGACAGTTCCTATCAAAAAGGATTTCAGAAAACATATACCATCTGTG 363
 QY 361 AAGAAAGTGGCCCTTCTCCCGCTTGCAAAATAGACATTTCAAAATTCGAAATGCGCAGC 420
 Db 364 AAGAAAGTGGCCCTTCTCCCGCTTGCAAAATAGACATTTCAAAATTCGAAATGCGCAGC 423
 QY 421 CAAGACCCCAATTTACCTGAAAGAGCCATATACAGAAAGGAAAGAAATTTAAACTGAG 480
 Db 424 CAAGACCCCAATTTACCTGAAAGAGCCATATACAGAAAGGAAAGAAATTTAAACTGAG 483
 QY 481 GGACATTCGTCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 Db 484 GGACATTCGTCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
 QY 541 CATTGGCAAGAGGGCCAGCAGATGTCTTTGGAGATATTTCTTTCTTCAAGGGAACATA 600
 Db 544 CATTGGCAAGAGGGCCAGCAGATGTCTTTGGAGATATTTCTTTCTTCAAGGGAACATA 603
 QY 601 CGAGCTTTTACCTGGAACAGGAGAAAGCACACTGGCCAGTTCCTCGGCGATAATGA 660
 Db 604 CGAGCTTTTACCTGGAACAGGAGAAAGCACACTGGCCAGTTCCTCGGCGATAATGA 663
 QY 661 GTTCTTCGGGGCAACAGCACCTCGGACTCTGTGTTTCACAGAAAGGCGCTCCCGGTGCT 720
 Db 664 GTTCTTCGGGGCAACAGCACCTCGGACTCTGTGTTTCACAGAAAGGCGCTCCCGGTGCT 723
 QY 721 CAAAATGCGCATCTCCCTCCGAGCATTTGGAGGATCCCAAGCTCTCATGTTGCGCTTATT 780
 Db 724 CAAAATGCGCATCTCCCTCCGAGCA-TGGAGGATCCCAAGCTCTCATGTTGCGCTTATT 782
 QY 781 GTCACCACTGACATTTAATCCAAACAGGAGTCTTCGGGCCACGA 826
 Db 783 GTCACCACTGACATTTAATCCAAACAGGAGTCTTCGGGCCACGA 826

RESULT 13

BQ680912

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 869)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2423 row: b column: 13

High quality sequence stop: 719.

Location/Qualifiers

1..869

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6261804"

FEATURES

source


```

/clone_lib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

```

BASE COUNT      255 a   134 c   166 g   275 t   19 others
ORIGIN
Query Match      30.8%; Score 783; DB 14; Length 869;
Best Local Similarity 96.5%; Pred. No. 3.2e-149;
Matches 802; Conservative 0; Mismatches 26; Indels 3; Gaps 1;

Qy 1042 ATGGAGCTCATCAAGGAAGACTAAGTCAGAGAGTCCCTCTCTGACCTTACAGGTTTC 1101
Db 1 ATGGAGCTCATCAAGGAAGACTAAGTCAGAGAGTCCCTCTCTGACCTTACAGGTTTC 60
Qy 1102 CCTCTCTCCTCGAGCTTGATCTGGCCCTCACCTTTGGATGAGTGTGAATGAAT 1161
Db 61 CCTCTCTCCTCGAGCTTGATCTGGCCCTCACCTTTGGATGAGTGTGAATGAAT 120
Qy 1162 GGATAAAATAAGTAACAAGATGCCAACTTTTTCCTTTGGGTAAAGTACAAAACA 1221
Db 121 GGATAAAATAAGTAACAAGATGCCAACTTTTTCCTTTGGGTAAAGTACAAAACA 180
Qy 1222 AACTAACACAGTGAAGAGAGGCTCCGAGAGTATTTTTCAGTTTGTCTGGTT 1281
Db 181 AACTAACACAGTGAAGAGAGGCTCCGAGAGTATTTTTCAGTTTGTCTGGTT 240
Qy 1282 TTCTAAATAATATTTTACAAAGTATTTTTCACCTGTTATGCCCTGTTTGCAAAACA 1341
Db 241 TTCTAAATAATATTTTACAAAGTATTTTTCACCTGTTATGCCCTGTTTGCAAAACA 300
Qy 1342 ATTTAGAAAAACAAACAGCAAACTATCTTGGCAAAAAGGAGTGTCTGAGC 1401
Db 301 ATTTAGAAAAACAAACAGCAAACTATCTTGGCAAAAAGGAGTGTCTGAGC 360
Qy 1402 CCATTTTCAGGAGCATTTGGTGTGTCGGCTCACATATTTTTCAGACACACAGAAA 1461
Db 361 CCATTTTCAGGAGCATTTGGTGTGTCGGCTCACATATTTTTCAGACACACAGAAA 420
Qy 1462 TCTGGCTTGGCCAGATTGGCACTAGCTATGAAGGCTGAGCGAGTCAATTAAGGAAT 1521
Db 421 TCTGGCTTGGCCAGATTGGCACTAGCTATGAAGGCTGAGCGAGTCAATTAAGGAAT 480
Qy 1522 TCAGGAACTTTATAGACTCCGACATTTCTGAGCAAGAGGAAAGTCAAAATTTATTTAA 1581
Db 481 TCAGGAACTTTATAGACTCCGACATTTCTGAGCAAGAGGAAAGTCAAAATTTATTTAA 540
Qy 1582 CACCTAAGCCCTTTTGTAGACTCTTTCTATATATTTGCTTAGGCTCACCATTGAAATTC 1641
Db 541 CACCTAAGCCCTTTTGTAGACTCTTTCTATATATTTGCTTAGGCTCACCATTGAAATTC 600
Qy 1642 TCCAGTGTAAACCTTTTCTGTTTCACATTTGAATTTATGGGTTTGGGGATTTCTT 1701
Db 601 TCCAGTGTAAACCTTTTCTGTTTTCACATTTGAATTTATGGGTTTGGGGATTTCTT 660
Qy 1702 GTAGTTCCTATATATCCCTATATATATATCTATATTCGAAATTTTGAAGTGTACGTAC 1761
Db 661 GTAGTTCCTATATATCCCTATATATATATCTATATTCGAAATTTTGAAGTGTACGTAC 720
Qy 1762 ATGTTGGTAAGACACAGCAAGTATTAAGTATTAAGTATTTTAAAGTATTAAGTAT 1821
Db 721 ATGTTGGTAAGACACAGCAAGTATTAAGTATTAAGTATTTTAAAGTATTAAGTAT 780
Qy 1822 ATTTT---ACGTGCTTTGGCTTTTATTCAGAGTCTACATTTTATAGA 1869
Db 781 ATTTT---ACGTGCTTTGGCTTTTATTCAGAGTCTACATTTTATAGA 1869

```

RESULT 14

```

BI915037      786 bp      mRNA      linear      EST 16-OCT-2001
LOCUS      603177226F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5241534 5',
DEFINITION      mRNA sequence.
ACCESSION      BI915037
VERSION      BI915037.1 GI:16179130
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE      1 (bases 1 to 786)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue procurement: Life Technologies, Inc.
              cDNA Library Preparation: by the I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov

```

```

Place: LLNL1609 row: c column: 07
High quality sequence stop: 786.
FEATURES
Location/Qualifiers
1..786
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:5241534"
    /clone_lib="NIH_MGC_121"
    /lab_host="DH10B"
    /note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
    Site: 2: EcoRV (destroyed); RNA source anonymous pool of 3
    fetal brains, female age 20 weeks, female age 24 weeks,
    and male age 26 weeks. Library is oligo-dT primed and
    directionally cloned (EcoRV site is destroyed upon
    cloning). Average insert size 1.7 kb, insert size range
    0.7-3.5 kb. Library is normalized and enriched for
    full-length clones and was constructed by C. Gruber
    (Invitrogen). Research Genetics tracking code 017. Note:
    this is a NIH_MGC Library."
BASE COUNT      195 a   225 c   192 g   174 t
ORIGIN
Query Match      30.4%; Score 774; DB 13; Length 786;
Best Local Similarity 99.9%; Pred. No. 2.2e-147;
Matches 785; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

Qy 19 CCGGCCACCGGCTGGGGCGCTTCACGCCGCTTCGAGCGGATATCGGTGAGCAGC 78
Db 1 CCGGCCACCGGCTGGGGCGCTTCACGCCGCTTCGAGCGGATATCGGTGAGCAGC 60
Qy 79 ACCACGCGCGGACGACTCGGTGATCTGCGACAGCGCAGGAGTTCGTCGCCGCCGCG 138
Db 61 ACCACGCGCGGACGACTCGGTGATCTGCGACAGCGCAGGAGTTCGTCGCCGCCGCG 120
Qy 139 GAGGCCGCGGACGAGCGGTGGGATCTCAGCGCGCGCGGTTTGTCTGTTGTTGTCAA 198
Db 121 GAGGCCGCGGACGAGCGGTGGGATCTCAGCGCGCGCGGTTTGTCTGTTGTTGTCAA 180
Qy 199 GACTGGATGATGAACCTGGCTCTAGGAAGCCTCACTTGGCCGTAACTCAGAAAGGTT 258
Db 181 GACTGGATGATGAACCTGGCTCTAGGAAGCCTCACTTGGCCGTAACTCAGAAAGGTT 240
Qy 259 CTCTTTGACCCCTCTCATTTTCGAAGCCACTTCTGAAGCCACTTGAAGCCACTTGA 318
Db 241 CTCTTTGACCCCTCTCATTTTCGAAGCCACTTCTGAAGCCACTTGAAGCCACTTGA 300

```


Db	824	CTCTGGAC	TTTACGGG	TTCCCTC	CTCCTCGAG	CTTGGATC	CTGGGCC	TTCACTTT	883
Qy	1141	GGA-	-TCAGT	CTGCTGA	TGTAATG	GATATAA	AAATAAG	TAAACAG	NTGCCA
Db	884	GGA	TTGAGG	CTTAAAT	GGATGG	ATAAATA	AGTAAAC	CAAGATG	GGCCAC
Qy	1199	TTGGG	GTAAAA	1209					
Db	944	TTTGG	GGGTAA	954					

Search completed: July 25, 2003, 19:53:57
Job time : 3785.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 15:39:40 ; Search time 563.197 seconds
(without alignments)
9326.039 Million cell updates/sec

Title: US-09-806-276A-4
Perfect score: 2546
Sequence: 1 cggggctagccggagacc.....cagctanaacgagggcagta 2546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues
Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1372	53.9	1832	15	US-10-153-668-16
c	441.4	17.3	486	11	US-09-796-692-7465
c	441.4	17.3	486	15	US-10-040-862-7465
c	374.6	14.7	445	12	US-09-918-995-5286
c	301	11.8	301	11	US-09-796-692-7516
c	301	11.8	301	15	US-10-040-862-7516
7	90.4	3.6	177	11	US-09-783-590-9071
8	85	3.3	413	10	US-09-864-761-26628
9	85	3.3	467	10	US-09-864-761-9983
10	82.4	3.2	169	11	US-09-783-590-6760
11	81.6	3.2	145	11	US-09-783-590-5334
12	62	2.4	11812	15	US-10-239-676-210
13	56.8	2.2	9539	15	US-10-239-676-52
c	53.6	2.1	416	11	US-09-960-352-4584
15	52.4	2.1	516	11	US-09-960-352-5785
16	52.2	2.1	446	11	US-09-960-352-3400

17	52.2	2.1	529	11	US-09-983-965-2109
18	51.6	2.0	9539	15	US-10-239-676-51
c	50.6	2.0	525	15	US-10-198-848-1483
20	49.8	2.0	344	11	US-09-960-352-1036
21	49.4	1.9	393	11	US-09-960-352-4582
22	49.2	1.9	11836	15	US-10-239-676-102
23	48.8	1.9	13606	15	US-10-239-676-165
24	48.4	1.9	8093	15	US-10-172-086-25
c	48.2	1.9	411	11	US-09-960-352-14521
26	48.2	1.9	480	11	US-09-960-352-5301
27	48	1.9	15732	15	US-10-239-676-96
28	47.6	1.9	431	11	US-09-960-352-5558
29	47.4	1.9	10619	15	US-10-239-676-1
30	47.2	1.9	8333	15	US-10-239-676-113
31	47.2	1.9	15732	15	US-10-239-676-95
32	47	1.8	6046	15	US-10-239-676-16
33	46.4	1.8	11036	15	US-10-239-676-118
c	46.2	1.8	5917	15	US-10-087-464-9
35	46.2	1.8	12968	15	US-10-239-676-202
36	46.2	1.8	14147	15	US-10-172-086-52
37	46	1.8	5689	15	US-10-239-676-90
38	46	1.8	6352	15	US-10-172-086-24
39	46	1.8	11260	15	US-10-239-676-19
40	45.8	1.8	6306	15	US-10-239-676-224
41	45.8	1.8	6944	15	US-10-172-086-112
42	45.6	1.8	7657	15	US-10-239-676-185
43	45.6	1.8	513509	12	US-09-754-853A-4
44	45	1.8	11036	15	US-10-239-676-117
45	44.8	1.8	7823	15	US-10-239-676-198

ALIGNMENTS

RESULT 1

US-10-153-668-16
; Sequence 16, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (508)..(837)
US-10-153-668-16

Query Match 53.9%; Score 1372; DB 15; Length 1832;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1174	GTAAAGATGCCAACTTTTTCCTTTGGGTAAAGGTACAAAAACAACCTAACCCACAG	1233
DB			
DB	1	GTAAAGATGCCAACTTTTTCCTTTGGGTAAAGGTACAAAAACAACCTAACCCACAG	60
QY			
DB	1234	TTGAAGAGAAGGGCTCCGGAGCTGATTTGCAGTGTTCGTGTGGTTCCTAAATAAT	1293
DB			
DB	61	TTGAAGAGAAGGGCTCCGGAGCTGATTTGCAGTGTTCGTGTGGTTCCTAAATAAT	120
QY			
DB	1294	ATTCTTACAAAGTATTTTTTACCTGGTATGCCCTGTTGCAGAAACAATTTAGAAAAA	1353
DB			
DB	121	ATTCTTACAAAGTATTTTTTACCTGGTATGCCCTGTTGCAGAAACAATTTAGAAAAA	180
QY			
DB	1354	ACACAAAGCAAAACCTATCTTTGGCAAAAAAGGAAGTGAGTCAGAGCCCATTTTCAGGA	1413
DB			
DB	181	ACACAAAGCAAAACCTATCTTTGGCAAAAAAGGAAGTGAGTCAGAGCCCATTTTCAGGA	240
QY			
DB	1414	GGCATTTGGTGATTTTCGGCTCACATATTTTGGAGACACACAAGAAATCTGGCTGGCC	1473
DB			
DB	241	GGCATTTGGTGATTTTCGGCTCACATATTTTGGAGACACACAAGAAATCTGGCTGGCC	300
QY			
DB	1474	AGGATTTGGCATAGCTATGAAGGGCTGAGCGAGTCACATTAAGSAACTTCACGGACATTT	1533
DB			
DB	301	AGGATTTGGCATAGCTATGAAGGGCTGAGCGAGTCACATTAAGSAACTTCACGGACATTT	360
QY			
DB	1534	ATAGCACTCCGACATTTCTGAGCAAGAGGAAGTCAAAATTTATTTAACACCTTAAGCCTT	1593
DB			
DB	361	ATAGCACTCCGACATTTCTGAGCAAGAGGAAGTCAAAATTTATTTAACACCTTAAGCCTT	420
QY			
DB	1594	TTTGTAGACTCTTTCTATATATTTAGCTTACCATAGCGAATTTCCAGTGTAAA	1653
DB			
DB	421	TTTGTAGACTCTTTCTATATATTTAGCTTACCATAGCGAATTTCCAGTGTAAA	480
QY			
DB	1654	ACTTTTCTGTTTCCACATTTCACTTTATGGGTTTTGGGAATTTCTTGTAGTCTTATA	1713
DB			
DB	481	ACTTTTCTGTTTCCACATTTCACTTTATGGGTTTTGGGAATTTCTTGTAGTCTTATA	540
QY			
DB	1714	TATCCCTATATATATCTATATTCGAAATTTTGACCTGCAGCTACATGTTGGTAGA	1773
DB			
DB	541	TATCCCTATATATATCTATATTCGAAATTTTGACCTGCAGCTACATGTTGGTAGA	600
QY			
DB	1774	CACAGGCAAGTATTAAGTAACTAGTTATTTTAAAGTAAATATATTTTACGTGC	1833
DB			
DB	601	CACAGGCAAGTATTAAGTAACTAGTTATTTTAAAGTAAATATATTTTACGTGC	660
QY			
DB	1834	CTTTGGCTTTTATTCGAGAGTACATTTTATAGATTCACACAGATGTGTCACCTA	1893
DB			
DB	661	CTTTGGCTTTTATTCGAGAGTACATTTTATAGATTCACACAGATGTGTCACCTA	720
QY			
DB	1894	TTTCCATTGGGATTCATTTGAAGCTGTATGTCGGTGTGGAAAGGTATTTATATAC	1953
DB			
DB	721	TTTCCATTGGGATTCATTTGAAGCTGTATGTCGGTGTGGAAAGGTATTTATATAC	780
QY			
DB	1954	TTAGTTTTTTTTTCTTCATCTGTTATCATACCTTTTAAACGCAACCAATTAACGGATTGA	2013
DB			
DB	781	TTAGTTTTTTTTTCTTCATCTGTTATCATACCTTTTAAACGCAACCAATTAACGGATTGA	840
QY			
DB	2014	AGTGAAGGCACAGGTTACTCATGATGCTTCTGCAGAGACTGTGGGCTACACCATAT	2073
DB			
DB	841	AGTGAAGGCACAGGTTACTCATGATGCTTCTGCAGAGACTGTGGGCTACACCATAT	900
QY			
DB	2074	GTTATTTGGAATATAGGTATTTTAGTACAGTACATGCTGCTTTATAGTAAATAACCAAGC	2133
DB			
DB	901	GTTATTTGGAATATAGGTATTTTAGTACAGTACATGCTGCTTTATAGTAAATAACCAAGC	960
QY			
DB	2134	GCACACATAAAAGTAAATGATAAAGTGAACCTGCTTTTATAGTAAATAACCAAGC	2193
DB			
DB	961	GCACACATAAAAGTAAATGATAAAGTGAACCTGCTTTTATAGTAAATAACCAAGC	1020
QY			
DB	2194	CATAAGAAATAAGTATAGCTAGAGAAATTTGCTTCTCTGAAATGTACATGAGCCCTTAAG	2253
DB			
DB	1021	CATAAGAAATAAGTATAGCTAGAGAAATTTGCTTCTCTGAAATGTACATGAGCCCTTAAG	1080

RESULT 2

RESULT 2
US-09-796-692-7465/c
; Sequence 7465, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES

```
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (395)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (415)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (436)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (446)
OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7465

Query Match 17.3%; Score 441.4; DB 11; Length 486;
Best Local Similarity 97.2%; Pred. No. 1.5e-104;
Matches 445; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1000 CCCGGACTGCCAGCGAGGACATGTTTACCATCCACCCCATCGAGCTCATCAAGG 1059
Db 459 CCCGGAGGTCCACCGAGGACANGTTTGACCATCCACCCCATNGAGCTCATCAAGG 400
Qy 1060 AAAGACTAAGTCAGAGGAGTCCCTCTCTGACCTTTACAGGTTCCCTCTCTCCCTGAGCT 1119
Db 399 AAAGNNTAAGTCAGAGGAGTCCCTCTCTGACCTTTACAGGTTCCCTCTCTCCCTGAGCT 340
Qy 1120 TGATCTTGGGCCCTCAGCTTTTGGATGAGGTCCTGAATGCTAATGATAAATAAGTAACA 1179
Db 339 TGATCTTGGGCCCTCAGCTTTTGGATGAGGTCCTGAATGCTAATGATAAATAAGTAACA 280
Qy 1180 AGATGCCAACCTTTTCCCTTTGGGTAAGGTACAAACAACTAACACAGTTGAAG 1239
Db 279 AGATGCCANCTTTTCCCTTTGGGTAAGGTACAAACAACTAACACAGTTGAAG 220
Qy 1240 AAGAGGCTTCCGGAGCTGTATTGCGAGTTTGTGTTGGTCTTTCTAAATAATATCTT 1299
Db 219 AAGAGGCTTCCGGAGCTGTATTGCGAGTTTGTGTTGGTCTTTCTAAATAATATCTT 160
Qy 1300 ACAAGATATTTTCTGCTGTTATGCGCTGTTTGGCAAAACAACTTTAGAAAAACACA 1359
Db 159 ACAAGATATTTTCTGCTGTTATGCGCTGTTTGGCAAAACAACTTTAGAAAAACACA 100
Qy 1360 AAGCAAAACCTATCTTGGCAAAAGAGTACAGTACAGCCCATTTTCAGGAGGCAAT 1419
Db 99 AAGCAAAACCTATCTTGGCAAAAGAGTACAGTACAGCCCATTTTCAGGAGGCAAT 40
Qy 1420 GGTGATGTTGGCTCACATATGTTTGCAGACACAA 1457
Db 39 GGTGATGTTGGCTCACATATGTTTGCAGACACAA 2

RESULT 3
US-10-040-862-7465/c
Sequence 7465, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040.862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186.126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190.479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200.545
PRIOR FILING DATE: 2000-04-27

OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (395)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (415)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (436)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (446)
OTHER INFORMATION: n=A,T,C or G
US-10-040-862-7465

Query Match 17.3%; Score 441.4; DB 15; Length 486;
Best Local Similarity 97.2%; Pred. No. 1.5e-104;
Matches 445; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1000 CCCGGACTGCCAGCGAGGACATGTTTACCATCCACCCCATCGAGCTCATCAAGG 1059
Db 459 CCCGGAGGTCCACCGAGGACANGTTTGACCATCCACCCCATNGAGCTCATCAAGG 400
Qy 1060 AAAGACTAAGTCAGAGGAGTCCCTCTCTGACCTTTACAGGTTCCCTCTCTCCCTGAGCT 1119
Db 399 AAAGNNTAAGTCAGAGGAGTCCCTCTCTGACCTTTACAGGTTCCCTCTCTCCCTGAGCT 340
Qy 1120 TGATCTTGGGCCCTCAGCTTTTGGATGAGGTCCTGAATGCTAATGATAAATAAGTAACA 1179
Db 339 TGATCTTGGGCCCTCAGCTTTTGGATGAGGTCCTGAATGCTAATGATAAATAAGTAACA 280
Qy 1180 AGATGCCAACCTTTTCCCTTTGGGTAAGGTACAAACAACTAACACAGTTGAAG 1239
```

Db 279 AGATGCCANCTTTTTCCTTTGGGTAAAGGTCACAAAACAACTAACACACAGTGAAG 220
QY 1240 AGAAGGGCTCCGGAGCTGATTTTGCAGTTTGTGGTGGTTCCTAAATAATATCTT 1299
Db 219 AGAAGGGCTCCGGAGCTGATTTTGCAGTTTGTGGTGGTTCCTAAATAATATCTT 160
QY 1300 ACAAGATATTTTACCTGTTATGCCCTGTTTGCACAAAACAACTTAGAAAAACAACA 1359
Db 159 ACAAGATATTTTACCTGTTATGCCCTGTTTGCACAAAACAACTTAGAAAAACAACA 100
QY 1360 AAGCAAAACCTATCTTGGCAAAAAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1419
Db 99 AAGCAAAACCTATCTTGGCAAAAAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 40
QY 1420 GGTGATGTTGGCTGCACATATTTTGCACACACAA 1457
Db 39 GGTGATGTTGGCTGCACATATTTTGCACACACAA 2

RESULT 4

US-09-918-995-5286
; Sequence 5286, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5286
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(445)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-5286

Query Match 14.7%; Score 374.6; DB 12; Length 445;
Best Local Similarity 94.2%; Pred. No. 3.7e-87;
Matches 389; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 CGGGGCTAGCCGGAGACCCGGCCACCGGCTGGGGGCGCTTCAGCGCGTCTCGGAGCGG 60
Db 31 CGGGGCTAGCCGGAGACCCGGCCACCGGCTGGGGGCGCTTCAGCGCGTCTCGGAGCGG 90
QY 61 ATAATGCGGTGAGCAGGACACCGCCGCGAGACTCGGCTGGTCTGCGCACAGCGGAGG 120
Db 91 ATAATGCGGTGAGCAGGACACCGCCGCGAGACTCGGCTGGTCTGCGCACAGCGGAGG 150
QY 121 GATTGCGTGGCGCCCGGAGCGCGGCGGAGCGGCTGGGATCTCAGCGCGCGCGGTT 180
Db 151 GATTGCGTGGCGCCCGGAGCGCGGCGGAGCGGCTGGGATCTCAGCGCGCGCGGTT 210
QY 181 TGCTCTGTTGTTGTTCAAGACTGGATGATGTAAGTGGCTCTCTAGGAAGCTCACTTGGC 240
Db 211 TGCTCTGTTGTTGTTCAAGACTGGATGATGTAAGTGGCTCTCTAGGAAGCTCACTTGGC 270
QY 241 GGTAACTTCAGGAAGGTTCTCTTGGACCCATCTCATTTGAGCCCACTTCTGAAGCCAC 300
Db 271 GGTAACTTCAGGAAGGTTCTCTTGGACCCATCTCATTTGAGCCCACTTCTGAAGCCAC 330
QY 301 TTGAGAAAATGATGTGACAGTTCCTATCAAAAAGAGTTTCAGAAACATATACCATCTGTG 360
Db 331 TTGAGAAAATGATGTGACAGTTCCTATCAAAAAGAGTTTCAGAAACATATACCATCTGTG 390

QY 361 AAGAAAGTGGCCCTTTCTCCCGTTTGCAAAATAGACATTTCAAAATCCAAAA 413
Db 391 AAGAAAGTGGCCCTTTCTCCCGTTTGCAAAATAGACATTTCAAAATCCAAAA 443

RESULT 5

US-09-796-692-7516/c
; Sequence 7516, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7516
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7516

Query Match 11.8%; Score 301; DB 11; Length 301;
Best Local Similarity 100.0%; Pred. No. 4.3e-68;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 709 CTCGCCGGTCTCAAAATGCCATCTCCCTCCGACCATTTGGAGGATCCCAAGCTCTCAT 768
Db 301 CTCGCCGGTCTCAAAATGCCATCTCCCTCCGACCATTTGGAGGATCCCAAGCTCTCAT 242
QY 769 GTTGCCCTTATTGTCACCAAGTGAATTAATCCAAACAGGAGTCTTCGGGCCAGCAAA 828
Db 241 GTTGCCCTTATTGTCACCAAGTGAATTAATCCAAACAGGAGTCTTCGGGCCAGCAAA 182
QY 829 GCTGCCAGGCTTAGTCTGGAGCCCGTCTATGGAGGAAAAGCTCAGGAGAAAAGCAGTCT 888
Db 181 GCTGCCAGGCTTAGTCTGGAGCCCGTCTATGGAGGAAAAGCTCAGGAGAAAAGCAGTCT 122
QY 889 GTTGAGAAATGGACAGTCCACAGGAGACACCTTCGTGGGCTCCAGGGTTCGTGATC 948
Db 121 GTTGAGAAATGGACAGTCCACAGGAGACACCTTCGTGGGCTCCAGGGTTCGTGATC 62
QY 949 TCAGTCCAGCCAGGACAGACAGCAGTCTCTCCAGCCTGTCCGAAACAGTACCCCGACTG 1008
Db 61 TCAGTCCAGCCAGGACAGACAGCAGTCTCTCTCCAGCCTGTCCGAAACAGTACCCCGACTG 2


```

; NAME/KEY: misc feature
; LOCATION: (122)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (128)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-5334

Query Match
Best Local Similarity 89.1%; Score 81.6; DB 11; Length 145;
Matches 106; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 966 GAGACAGCACTCCCTCCAGCCTGTCGGAACAGTACCCCGACTGCGCCAGCCGAGACATGT 1025
DB 1 GGNAGAGCCACTCCCTCCAGCCTNTCCGANCAGTACCCCGACTGCGCCAGCCGAGACATGT 60
QY 1026 TTGACATCCCAACCCCATCGCG-AGCTCATCAAGGG-AAAGACTAAGTCAGAGGAGTCC 1082
DB 61 TTNACCATCCCAACCCCATGTGAAGTNCANCAAGGNNAGACTAAGTCAGAGGAGTCC 119

RESULT 12
US-10-239-676-210
; Sequence 210, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 210
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-210

Query Match
Best Local Similarity 47.2%; Score 62; DB 15; Length 11812;
Matches 286; Conservative 0; Mismatches 315; Indels 5; Gaps 3;

QY 1570 AAATTATTATTAACACTAAGCCTTTTGTAGACTCTTTTCTATATATATGCTTAGGCAC 1629
DB 2370 AAATTATTATTAGTTTAAATATATATAAAATAGGTGTTATTTGATTTTAGGTATA 2429
QY 1630 CATACGGAATCTCCAGCTGTTAAACTTCTCTGTTTTCACATTTGAACTTTATCGGTTT 1689
DB 2430 GTTTTGTGATCGTC--GTTTAAATTTTGTGAAATATATATTTTAAATAGATTA 2487
QY 1690 GGGGATTTCTTGTAGTCTTATATATATATATATATATATATATATATATATATAT 1749
DB 2488 GAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 2547
QY 1750 ACTGTCAGCATCATGTTGGTAAGACACAGCAAGCAAGTATTACTGTAAGTAACTTATTTTA 1809
DB 2548 ATGATTTTAAAGCTTTGAGTTTAAATTTTATTTTAAATTAATGATAATTTT 2607
QY 1810 AAGTTAAATATATTTTACGCGCTTTTGTGCTTTTATATGAGAGTCTACATTTTATAGA 1869

; NAME/KEY: misc feature
; LOCATION: (122)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (128)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-5334

Query Match
Best Local Similarity 89.1%; Score 81.6; DB 11; Length 145;
Matches 106; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 966 GAGACAGCACTCCCTCCAGCCTGTCGGAACAGTACCCCGACTGCGCCAGCCGAGACATGT 1025
DB 1 GGNAGAGCCACTCCCTCCAGCCTNTCCGANCAGTACCCCGACTGCGCCAGCCGAGACATGT 60
QY 1026 TTGACATCCCAACCCCATCGCG-AGCTCATCAAGGG-AAAGACTAAGTCAGAGGAGTCC 1082
DB 61 TTNACCATCCCAACCCCATGTGAAGTNCANCAAGGNNAGACTAAGTCAGAGGAGTCC 119

RESULT 13
US-10-239-676-52
; Sequence 52, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 52
; LENGTH: 9539
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-52

Query Match
Best Local Similarity 46.1%; Score 56.8; DB 15; Length 9539;
Matches 190; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

QY 1577 TTTAACACCTTAAGCCCTTTTGTAGACTCTTTTCTATATATGCTTAGGCCTCACCACG 1636
DB 21 TTTAGATATAGAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 80
QY 1637 AATCTCCAGTGTAAACTTTTCTGTTTTCACATTTGAACCTTTATGCGTTTGGGAT 1696
DB 81 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 140
QY 1697 TTTTGTAGTCTTATATATATATATATATATATATATATATATATATATATATATAT 1756

```


us-09-806-276a-4.rni

Tue Jul 29 09:34:16 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 15:25:15 ; Search time 123.391 Seconds
(without alignments)
6328.365 Million cell updates/sec

Title: US-09-806-276A-4
Perfect score: 2546
Sequence: 1 cgggctagcccgagagacc.....cagctanaacagggcagta 2546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	52	2.0	19124	2	US-08-487-826B-13
c 2	45.4	1.8	7218	1	US-08-232-463-14
c 3	44	1.7	6243	2	US-09-056-075-1
c 4	42.8	1.7	19124	2	US-08-487-826B-13
c 5	41	1.6	8220	2	US-08-568-459A-11
c 6	41	1.6	8220	2	US-08-487-826B-11
c 7	41	1.6	8220	4	US-09-210-288-11
c 8	40	1.6	2264	3	US-08-471-733-5
c 9	40	1.6	2264	3	US-08-468-878-5
c 10	40	1.6	2264	4	US-08-750-494-5
c 11	40	1.6	660	1	US-07-991-867B-32
c 12	39.2	1.5	660	1	US-08-107-755A-32
c 13	39.2	1.5	660	2	US-08-544-332-32
c 14	39.2	1.5	660	2	US-09-370-861A-32
c 15	39.2	1.5	660	4	US-08-998-416-534
c 16	39.2	1.5	854	1	US-07-991-867B-8
c 17	39.2	1.5	1511	1	US-08-107-755A-8
c 18	39.2	1.5	1511	2	US-08-544-332-8
c 19	39.2	1.5	1511	4	US-09-370-861A-8
c 20	39.2	1.5	1511	4	US-08-852-629-11
c 21	39.2	1.5	4810	3	US-08-852-629-15
c 22	39.2	1.5	4838	3	US-08-998-416-595
c 23	39	1.5	658	4	US-08-998-416-595
c 24	38.8	1.5	473	1	US-08-764-100-16
c 25	38.8	1.5	821	4	US-08-998-416-541
c 26	38.8	1.5	1483	4	US-08-961-527-315
c 27	38.8	1.5	4970	1	US-08-764-100-14

28	38.8	1.5	4970	1	US-08-764-100-20
c 29	38.6	1.5	6095	4	US-09-357-206A-18
c 30	38.6	1.5	6325	4	US-09-357-206A-20
c 31	38.6	1.5	6428	4	US-09-357-206A-22
c 32	38.6	1.5	7400	1	US-08-261-663A-1
c 33	38.6	1.5	7400	5	PCT-US95-07754A-1
c 34	38.6	1.5	12286	4	US-09-357-206A-1
c 35	38.6	1.5	38564	4	US-09-734-673-3
c 36	38	1.5	20674	4	US-09-641-638-651
c 37	37.4	1.5	3457	4	US-09-293-549-3
c 38	37.2	1.5	5852	1	US-07-867-106-2
c 39	37.2	1.5	11091	4	US-09-134-001C-2243
c 40	37.2	1.5	28001	4	US-09-819-993-3
c 41	36.8	1.4	1850	3	US-08-617-860B-32
c 42	36.8	1.4	2781	3	US-08-749-522-4
c 43	36.8	1.4	4098	2	US-08-605-106-4
c 44	36.6	1.4	1780	1	US-08-434-702-3
c 45	36.6	1.4	1780	1	US-08-271-883-3

ALIGNMENTS

RESULT 1

US-08-487-826B-13/c

Sequence 13, Application US/08487826B

Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH21.001CPI

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 19124 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-487-826B-13

Query Match

Best Local Similarity

2.0%; Score 52; DB 2; Length 19124;

46.9%; Pred. No. 0.0011;

```
;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pt2gpt-Fts
US-08-232-463-14

Query Match      1.8%; Score 45, 4; DB 1; Length 7218;
Best Local Similarity 5.7%; Pred. No. 0.039; Mismatches 163; Indels 0; Gaps 0;
Matches 22; Conservative 202;

QY   1573 TTTATTACACTTAGCCTTTTGTAGACTCTTTCTATATATATTCCTTAGCTGCACCAT 1632
Db    : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1095 YYYYYYYYYYYYYYYYYYYYVVVYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1154
QY   1633 AGCGAATCTCCAGTGTTAAACCTTTTCTGTTTCACATTTCACTTTATGGTTTTGGG 1692
Db    : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1155 YYYYYYYYYYYYYYYYYYYYVVVYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1214
QY   1693 GATTTCTGTGAGTCTTATATAATCCCATAFATTAATATCTATATGCAAAATTTGAC 1752
Db    : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1215 YYYYYYYYYYYYYYYYYYYYVVVYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1274
QY   1753 GTCAGCTACATGTTGGTAAGACACAGCAAAGTATTACTGTAAC TAGTTATTTTAAAG 1812
Db    : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1275 YYYYYYYYYYYYYYYYYYYYVVVYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1334
QY   1813 TTAAAATATATTTTACGTGCCCTTTTATTCGAGAGCTCACATTTTATATAGATTC 1872
Db    : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1335 YYYYYYYYYYYYYYYYYYYYVVVYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1394
QY   1873 TACAFCAGATGTTGTCACCTATTTCCATTTGGGATTCCAATGTAAGCTGTGAFTGCGGTG 1932
Db    : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1395 YYYYYYYYYYYYYYYYYYYYVVVYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1454
QY   1933 TTTCGAAAGCTGTTATCACTACTAGTT 1959
Db    ||| | | | | | | | | |
1455 TTAACTACTGTCATAGTAGTAAT 1481
```

1873	TACATCAGATGTTGTCACTATTTCATGGGATGCCAATGTAACTCGTGAATGTCGTG	1394
QY	: : : :	:
Dc	: : : :	:
1395	YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY	1454
QY	TTTTGGAAAAGTGTAATCATACTAGTT	1599
Dc		1481
1455	TTTAACTACTTGCATAGTAGGTAAT	

RESULT 3
US-09-056-075-1/c
; Sequence 1, Application US/09056075
; Patent No. 5955368
; GENERAL INFORMATION:
; APPLICANT: Johnson, Eric A.
; APPLICANT: Bradshaw, Marite
; APPLICANT: Rood, Julian
; TITLE OF INVENTION: Expression System for Clostridium
; TITLE OF INVENTION: Species
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; City: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,075
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95238
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000

us-09-806-276a-4.rni

Tue Jul 29 09:34:16 2003

TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6243 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc.feature
LOCATION: 3770..4013
OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orit) from
OTHER INFORMATION: plasmid RP4"
US-09-056-075-1

Query Match 1.7%; Score 44; DB 2; Length 6243;
Best Local Similarity 47.2%; Pred. No. 0.087;
Matches 134; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY 1706 TTCCTATATATCCCTATATATATATATATATATGCAAAATTTTGACTGTCAGTCAATGT 1765
Db 1487 TTATAAGGCGTCATTTTAT 1428
QY 1766 TGGTAAGACACAGCAAGATTAAGTAACTAAGTATATATATATATATATATATATAT 1825
Db 1427 TTTCGAACCTTTAAATAAAATAATATATATATATATATATATATATATATATATAT 1368
QY 1826 TTACGTGCTGCTGCTTTTATATGCAAGTCTACATTTTATAGATCTACATGATGT 1885
Db 1367 TTATTTTATTTTAT 1308
QY 1886 GTCACATATTTCCATGCGGATTCATGTAAGCTGCTGTAAGTGGGCTTTGGAAAGTGT 1945
Db 1307 TTTTATATTTTAT 1248
QY 1946 ATTCATATAGTATTTTATATATATATATATATATATATATATATATATATATAT 1989
Db 1247 TTATTTTAT 1204

RESULT 4
US-08-487-826B-13
Sequence 13, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 1.7%; Score 42.8; DB 2; Length 19124;
Best Local Similarity 51.2%; Pred. No. 0.34;
Matches 124; Conservative 0; Mismatches 117; Indels 1; Gaps 1;
QY 1732 CTATATGCAAAATTTTGACTGTCAGTCAATGTTGGTAAGACACAGCAAGATTAAT 1791
Db 4282 CAATTTGCAATATGTAATTAATAAATAATATATATAAAGAAATATATAAATAATAT 4341
QY 1792 GTAACCTAAGTATTTTAAAGTTAAATATATATATATATATATATATATATATAT 1851
Db 4342 ATAAAAAATTAAGCATAAATGTCACATAAATTTTATTTTATTAATTAATTTATTTT 4401
QY 1852 GAG-TCTACATTTTATAGATTTCTACATCAGATGTTGTCACATTTTCCATTTGGAT 1910
Db 4402 TTGTTCTAAAATATATGTAATGATTAAGAAATATATTTTGTGCTAATAATAAAGAT 4461
QY 1911 TTGTAAGCTGTATGTCGCGTGGTGGAAAGCTGATTAATCACTAGTTATTTTCTTC 1970
Db 4462 CTAATATTAATTTAT 4521
QY 1971 AT 1972
Db 4522 AT 4523

RESULT 5

US-08-568-459A-11
Sequence 11, Application US/08568459A
Patent No. 5849306

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,459A
FILING DATE: 07-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 8220 base pairs

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NHH121.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
TYPE: nucleic acid
LENGTH: 8220 base pairs
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum


```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 303..2162
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 303..365
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 366..2159
; US-08-471-733-5

Query Match
Best Local Similarity 1.6%; Score 40; DB 3; Length 2264;
Matches 130; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 1601 ACTCTTTTCTATATATGCTAGGCTCACCATAGGAATTCGCCAGTGTAAACCTTTTC 1660
Db 328 AATTATATAAATATGATTTTCATATATATACCTCTCTGTGTAACTATAAATTTTT 269
QY 1661 TGTTTTCAATTTGAACCTTTATGGGTTTGGGGAATTTCTTGTAGTCTTATATATCCCT 1720
Db 268 ATTAGTTTTAGTTATTTTAAACGCTCTTAATAATATATATAATACATATTTACTTATT 209
QY 1721 ATATATTATCTATATATGCTAGGCTCACCATAGGAATTCGCCAGTGTAAACCTTTTC 1780
Db 208 GTACATAATACATGTATACATATTTAAACCTATTAGATAAAATTTCAATATCATATTGG 149
QY 1781 AAAGTATTACTGTAAGTTATTTTAAAGTTAAATATATTTTACGTGCCTTTTGGC 1840
Db 148 TATTGTGATTTTATCTATTTCAAAATTTTAAATTAAGATTAAATTCATTGTCACTAATC 89
QY 1841 TTTTATTCGACAGTCTACATTTTATAGATTTCTACATCAG 1880
Db 88 TTATTTATTTAAATGATTAATTTTATACATAAATAAACACG 49

RESULT 10
US-08-468-878-5/C
Sequence 5, Application US/08468878
Patent No. 6090586
GENERAL INFORMATION:
APPLICANT: BERGSTROM SVEN
APPLICANT: BARBOUR ALAN G.
TITLE OF INVENTION: NEW 66 KDA ANTIGEN FROM BORRELIA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 SEVENTH STREET, N.W.
CITY: WASHINGTON
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,878
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/262,220
FILING DATE: 20-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: BERGSTROM-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 2264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Borrelia afzelii
; STRAIN: ACAI
; IMMEDIATE SOURCE:
; CLONE: PJB-104
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 303..2162
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 303..365
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 366..2159
; US-08-468-878-5

Query Match
Best Local Similarity 1.6%; Score 40; DB 3; Length 2264;
Matches 130; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 1601 ACTCTTTTCTATATATGCTAGGCTCACCATAGGAATTCGCCAGTGTAAACCTTTTC 1660
Db 328 AATTATATAAATATGATTTTCATATATATACCTCTCTGTGTAACTATAAATTTTT 269
QY 1661 TGTTTTCAATTTGAACCTTTATGGGTTTGGGGAATTTCTTGTAGTCTTATATATCCCT 1720
Db 268 ATTAGTTTTAGTTATTTTAAACGCTCTTAATAATATATATAATACATATTTACTTATT 209
QY 1721 ATATATTATCTATATATGCTAGGCTCACCATAGGAATTCGCCAGTGTAAACCTTTTC 1780
Db 208 GTACATAATACATGTATACATATTTAAACCTATTAGATAAAATTTCAATATCATATTGG 149
QY 1781 AAAGTATTACTGTAAGTTATTTTAAAGTTAAATATATTTTACGTGCCTTTTGGC 1840
Db 148 TATTGTGATTTTATCTATTTCAAAATTTTAAATTAAGATTAAATTCATTGTCACTAATC 89
QY 1841 TTTTATTCGACAGTCTACATTTTATAGATTTCTACATCAG 1880
Db 88 TTATTTATTTAAATGATTAATTTTATACATAAATAAACACG 49

RESULT 11
US-08-750-494-5/C
Sequence 5, Application US/08750494
Patent No. 6204018
GENERAL INFORMATION:
APPLICANT: BERGSTROM SVEN
APPLICANT: BARBOUR ALAN G.
TITLE OF INVENTION: NEW 66 KDA ANTIGEN FROM BORRELIA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 SEVENTH STREET, N.W.
CITY: WASHINGTON
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,494
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/262,220
FILING DATE:
```


[illegible]

Db 1741 AAAATTTTGACTGTCAGCTACATGTTGGTAAGACACAGGCAAAAGTATTACTGTAACTAAG 1800
Qy 1801 TTAATTTTAAAGTTAAATATATTTTACGTGCTTTGGCTTTTATTTGACAGAGTCAC 1860
Db 1801 TTAATTTTAAAGTTAAATATATTTTACGTGCTTTGGCTTTTATTTGACAGAGTCAC 1860
Qy 1861 TTTTATAGATCTACATCAGATGTTGTCACATTTATTTCCATTTGGGATTCGATGAGCTG 1920
Db 1861 TTTTATAGATCTACATCAGATGTTGTCACATTTATTTCCATTTGGGATTCGATGAGCTG 1920
Qy 1921 TGTATGTCGCTGTTGGAAAAGTATTCATACATTTAGTTTATTTTCTTCATCTGTTATC 1980
Db 1921 TGTATGTCGCTGTTGGAAAAGTATTCATACATTTAGTTTATTTTCTTCATCTGTTATC 1980
Qy 1981 ATACTTTTAAAGCAACCAATTAACGGATGTAAGAGTGAAGGACACAGGTTTACTCATGAT 2040
Db 1981 ATACTTTTAAAGCAACCAATTAACGGATGTAAGAGTGAAGGACACAGGTTTACTCATGAT 2040
Qy 2041 GCTTCTGAGAGACTGTGGCTACACACATATGTTATTTGGAATATAGTATTTTAGT 2100
Db 2041 GCTTCTGAGAGACTGTGGCTACACACATATGTTATTTGGAATATAGTATTTTAGT 2100
Qy 2101 ACAGTACATCTGCTGATTTACATGATGTTCAAGCAACACAAATAAAAGTAAATGATAAA 2160
Db 2101 ACAGTACATCTGCTGATTTACATGATGTTCAAGCAACACAAATAAAAGTAAATGATAAA 2160
Qy 2161 GTGAACCTGCTGTTTATAGTAAATAAACAAGCAACCAATAGAGATATAGCTAGAGAA 2220
Db 2161 GTGAACCTGCTGTTTATAGTAAATAAACAAGCAACCAATAGAGATATAGCTAGAGAA 2220
Qy 2221 ATTGCTTCTGCTGAAATGATACATGAGCCCTTAAGGTAAGAGATGATTTCCATCTACTCTCA 2280
Db 2221 ATTGCTTCTGCTGAAATGATACATGAGCCCTTAAGGTAAGAGATGATTTCCATCTACTCTCA 2280
Qy 2281 TTTTGATTAATCTCTGATGTTGAGAGCTAGAACTGAGCCCTCTCTACTTTTGGAAAA 2340
Db 2281 TTTTGATTAATCTCTGATGTTGAGAGCTAGAACTGAGCCCTCTCTACTTTTGGAAAA 2340
Qy 2341 ATGAACATGAGGTCAGATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2400
Db 2341 ATGAACATGAGGTCAGATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2400
Qy 2401 CAGTGTGCTATTTCTGAGCATCTTCTGACTTGAACACCTTCTACAGCAAACTTTGCAAG 2460
Db 2401 CAGTGTGCTATTTCTGAGCATCTTCTGACTTGAACACCTTCTACAGCAAACTTTGCAAG 2460
Qy 2461 TCCAGTTTCATCCCTGTAAGGCAATGCTTTTTCAGGCAAGAGTGCATATAGACGAGA 2520
Db 2461 TCCAGTTTCATCCCTGTAAGGCAATGCTTTTTCAGGCAAGAGTGCATATAGACGAGA 2520
Qy 2521 TAAAGGCAGCTAAGAGGGGAGTA 2546
Db 2521 TAAAGGCAGCTAAGAGGGGAGTA 2546

RESULT 2

ABK64131/C

ID ABK64131 standard; DNA; 564 BP.

XX

XX

18-JUN-2002 (first entry)

XX

Human benign prostatic hyperplasia gene #26.

Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.

Homo sapiens.

OS

PN

W0200212440-A2.

XX

14-FEB-2002.

XX

07-AUG-2001; 2001WO-US24708.
07-AUG-2000; 2000US-223323P.
05-JUN-2001; 2001US-0873319.
(GENE-) GENE LOGIC INC.
(NISB) JAPAN TOBACCO INC.
Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
WPI; 2002-257476/30.
Identifying drugs for and diagnosing benign prostatic hyperplasia, by detecting expression levels of one or more genes in prostate cells from patient that are differentially regulated compared to normal prostate cells
Disclosure; Page 83; 444pp; English.
The invention relates to a method of diagnosing (I) the onset or progression of benign prostatic hyperplasia (BPH), or screening (II) for identifying an agent that modulates the onset or progression of BPH. The method is based on changes in gene expression in BPH tissue isolated from patients exhibiting different clinical states of prostate hyperplasia as compared to normal prostate tissue. (I) comprises detecting the expression levels of one or more genes in prostate cells from the subject that are differentially regulated compared to normal prostate cells. (II) comprises preparing a first gene expression profile of BPH cells or BPH-like cell population, exposing the cells to the agent, preparing a second gene expression profile of the agent exposed cells, and comparing the first and second gene expression profiles. (I) is useful for diagnosing the onset or progression of BPH. (II) is useful for identifying an agent that modulates the onset or progression of BPH. The methods are useful to present information identifying the expression level in a tissue or cells, by comparing the expression level of genes given in the specification in the tissue or cells to the level of expression of gene in the database, and displaying the expression levels of at least one gene in the tissue or cell sample compared to the expression level in BPH. Agents using (II) are useful for treating BPH or prostate cancer. ABK64106-ABK64860 represent human benign prostatic hyperplasia gene sequences of the invention.

SQ Sequence 564 BP; 210 A; 93 C; 85 G; 174 T; 2 other;

Query Match 19.3%; Score 491.8; DB 24; Length 564;
Best Local Similarity 98.6%; Pred. No. 2.1e-118;
Matches 506; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Qy 1653 AACCTTTCTGTTTTCACATTTGAACCTTTATGGGTTTGGGATTTCTTCTAGTCTTAT 1712
Db 523 AAACCTTTCTGTTTTCACATTTGAACCTTTATGGGTTTGGGATTTCTTCTAGTCTTAT 464
Qy 1713 ATATCCCTAT 1772
Db 463 ATATCCCTAT 404
Qy 1773 ACACAGGCAAAAGTATTTACTGTAACCTTAAGTTATTTTAAAGTTAAATATATATTTTACGTG 1832
Db 403 ACACAGGCAAAAGTATTTACTGTAACCTTAAGTTATTTTAAAGTTAAATATATATTTTACGTG - 345
Qy 1833 CCTTTGGCTTTTATTTGACAGAGCTACATTTATAGATTTTACATCAGATCTCTCACTT 1892
Db 344 CCTTTGGCTTTTATTTGACAGAGCTACATTTATAGATTTTACATCAGATCTCTCACTT 285
Qy 1893 ATTTCCATTTGGGATTTCCATTTGTAAGCTGTATGTGCGTGTGTTGTTGAAAAGTGTATTTCATA 1952
Db 284 ATTTCCATTTGGGATTTCCATTTGTAAGCTGTATGTGCGTGTGTTGTTGAAAAGTGTATTTCATA 225
Qy 1953 CTTAGTCTTTTCTTCTCATCTGTTATCATCTTTTAAACAGACCAACCAATACGGATGTGA 2012
Db 224 CTTAGTCTTTTCTTCTCATCTGTTATCATCTTTTAAACAGACCAACCAATACGGATGTGA 165
Qy 2013 AAGTAAAGGCACAGGTTTACTCATGATGCTTCTGCAGAGACTGTGGGCTACACACATA 2072

Db 164 AAGTGTAAAGCACAGGTTACTCATGATGCTTCTGCAGACACTGTGGCTACACACATA 105
 QY 2073 TGTATTGGAATATAGGTATTTTAGTACAGTACATCTTGCATTACATAGTACTTCA 2132
 Db 104 TGTATTGGAATATAGGTATTTTAGTACAGTACATCTTGCATTACATAGTACTTCA 45
 QY 2133 AGCAACAATAAAGTAATAATGATAAAGTGAA 2165
 Db 44 AGCAACAATAAAGTAATAATGATAAAGTGAA 12

RESULT 3

AAS89293
 ID AAS89293 standard; cDNA; 1617 BP.
 AC
 AS89293;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #25097.

Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG25106.

New isolated polynucleotide and encoded polypeptides, useful in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits and to assess
 biodiversity -

Claim 1; SEQ ID No 25097; 103pp; English.

The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 and gene mapping, and in recombinant production of (II). The
 polynucleotides are also used in diagnostics as expressed sequence tags
 for identifying expressed genes. (I) is useful in gene therapy techniques
 to restore normal activity of (II) or to treat disease states involving
 (II). (II) is useful for generating antibodies against it, detecting or
 quantitating a polypeptide in tissue, as molecular weight markers and as
 a food supplement. (II) and its binding partners are useful in medical
 imaging of sites expressing (II). (I) and (II) are useful for treating
 disorders involving aberrant protein expression or biological activity.
 The polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 and to produce other types of data and products dependent on DNA and
 amino acid sequences. AAS64197-AAS94564 represent novel human
 diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed
 specification, but was obtained in electronic format directly from WIPO
 at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1617 BP; 329 A; 393 C; 461 G; 434 T; 0 other;

Query Match 13.0%; Score 330.8; DB 23; Length 1617;
 Best Local Similarity 90.5%; Pred. No. 5.3e-76;
 Matches 353; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
 QY 788 GTGACATTAAATTCACACAGAGTCCTTCGGCCAGCAAAAGCTGCCAGGCTTAGCTGC 847
 Db 1228 GTGGCTGTGACATGGAAAGAGTCCCTTCGGCCAGCAAAAGTTCGCAAGCTTAGCTGC 1287
 QY 848 GAGCCCGTCATGGAGGAAAAAGCTCAGGAGAAAAAGCAGTCTGTGGAGAAATGGACAGTC 907
 Db 1288 GAGCCCGTCATGGAGGAAAAAGCTCAGGAGAAAAAGCAGTCTGTGGAGAAATGGACAGTC 1347
 QY 908 CACCAGGAGACACCTCGTGGGGCTCAGCCGGTTCGCATCTCAGTCCAGCAGGAGGAGA 967
 Db 1348 CACCAGGAGACACCTCGTGGGGCTCAGCCGGTTCGCATCTCAGTCCAGCAGGAGGAGA 1407
 QY 968 GACAGCCACTCTCCAGCCTGTCGGAACAGTACCCGACTGGCCAGCCGAGGACATGTTT 1027
 Db 1408 GACAGCCACTCTCCAGCCTGTCGGAACAGTACCCGACTGGCCAGCCGAGGACATGTTT 1467
 QY 1028 GACATCCACCCCATGGAGCTCATCAAGGAAAAAGACTAAGTCAGAGGAGTCCCTCTCT 1087
 Db 1468 CACCATCCCGGTCATGCAAGCTCATCAAGGAAAAAGACTAAGTCAGAGGAGTCCCTCTCT 1527
 QY 1088 GACCTTACAGGTTCCCTCTCTCCCTCGCAGCTTGATCTGGCCCTCAGTCTTTGGATGAG 1147
 Db 1528 GATCTTACAGGTTCCCTCTCTCCCTCGCAGCTTGATCTGGCCCTCAGTCTTTGGATGAG 1177
 QY 1148 GTGCTGAATGTAATGGATAAAATAAGTAA 1177
 Db 1588 GGGCTGAATGTCACGGAAAAATAAGTAA 1617

RESULT 4

AAS94524
 ID AAS94524 standard; cDNA; 156 BP.
 AC
 AS94524;

18-JUL-2000 (first entry)

Cartilage-associated polynucleotide D3.

Chondrocyte; cartilage; differentiation; marker; human; ss.

Homo sapiens.

WO200017359-A1.

30-MAR-2000.

23-SEP-1999; 99WO-US22108.

24-SEP-1998; 98US-0101598.

(GENZ) GENZYME CORP.

Haudenschild DR, Binette F;

WPI; 2000-283584/24.

New isolated nucleic acid molecule corresponding to genes involved in
 the process of chondrocyte differentiation and dedifferentiation is
 useful for detecting a chondrocyte in a biological sample -

Example; Page 70; 75pp; English.

This is the nucleotide sequence of human cDNA clone D3, one of 99
 novel polynucleotides of the invention (see AAS94429-294527) that are
 involved in modulation of cartilage growth and differentiation. D3
 is deposited in Escherichia coli DH10B clone C55 as ATCC 98891. The
 99 polynucleotides were isolated from subtracted libraries during

Tue Jul 29 09:34:15 2003

articular chondrocyte differentiation. They correspond to portions of genes that are differentially regulated in human articular chondrocytes 48 hr after transfer to an alginate suspension culture as compared to those expressed in dedifferentiated monolayer culture. The polynucleotides can be used as markers to identify chondrocytes, as probes for detecting chondrocytes, or can be incorporated into vectors for recombinant production of chondrocyte differentiation-related polypeptides. Such polypeptides are useful for screening for modulator compounds and for production of antibodies used for in vivo imaging or immunoassay.

Sequence 156 BP; 37 A; 26 C; 27 G; 57 T; 9 other;

Query Match 5.3%; Score 133.8; DB 21; Length 156; Best Local Similarity 91.1%; Pred. No. 5.7e-25; Matches 144; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

2239 ACATGAGCCCTTAAGTAAGATGATTTCCATCTACTCTCATTTGATTACTTCTTAT 2298
1 ACATGAGCCCTTAAGTAAGATGATTTNNATCTACTCTCTTTTGTACNWCYCTAT 60

2299 GGTGTAGAGCTAGAACTGAGCCTCTCTACTTTTGGAAAAATGAACATGTAGGTCAG 2358
61 GGTGTAGAGCTAGAACTGAGCCTCTCTACTTTTGGAAAAATGAACATGTAGGTCAG 118

2359 ATTTTGTATTTTTTTTTTTTAAAGTCAGCACTGATGCCACC 2396
119 ATTTTGTATTTTTTTTTTTTAAAGTCAGCACTGATGCCACC 156

RESULT 5
ABQ17530
ID ABQ17530 standard; DNA; 1079 BP.
XX
XX ABQ17530;
AC
XX
XX 12-JUL-2002 (first entry)
DT
XX
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 4121.
DE
XX
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200218632-A2.
PN
XX
XX 07-MAR-2002.
PD
XX
XX 01-SEP-2001; 2001WO-EP10074.
PF
XX
XX 01-SEP-2000; 2000DE-1043826.
PR
XX
XX 05-SEP-2000; 2000DE-1044543.
XX
XX (EPIC-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
PI
XX
XX WPI; 2002-371829/40.
DR
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
PS
XX
XX This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic

DNA that contains the target C is amplified to form a labeled amplicon. The amplicon is hybridised to two classes, each with at least one member of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP's); and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention.

Sequence 1079 BP; 140 A; 124 C; 384 G; 431 T; 0 other;

Query Match 4.9%; Score 124.8; DB 24; Length 1079; Best Local Similarity 78.1%; Pred. No. 3.7e-22; Matches 150; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

1 CGGGGTAGCCCGGAGACCGCCGCGCTTACGGCGCTTACGGCGCTTACGGCGCGG 60
724 CGGGGTAGTTTGGAGATTTCGGTTATCGGTTTGGCGGTTTACGTCGTTTCGGAGCGG 783
61 ATATGCGGTGAGCAGGACCGCCGCGCAGACTCGGCTGGATCTCGCACACGCGCAGG 120
784 ATATGCGGTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 843
121 GATTGCGTGGCGCCCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTT 180
844 GATTGCGTGGCGGTT 903
181 TGTCTGTTGT 192
904 AGTCGGGTAGT 915

RESULT 6
ABQ17531/C
ID ABQ17531 standard; DNA; 1079 BP.
XX
XX ABQ17531;
AC
XX
XX 12-JUL-2002 (first entry)
DT
XX
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 4122.
DE
XX
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200218632-A2.
PN
XX
XX 07-MAR-2002.
PD
XX
XX 01-SEP-2001; 2001WO-EP10074.
PF
XX
XX 01-SEP-2000; 2000DE-1043826.
PR
XX
XX 05-SEP-2000; 2000DE-1044543.
XX
XX (EPIC-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
PI
XX
XX WPI; 2002-371829/40.
DR
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful
PT

us-09-806-276a-4.rng

Tue Jul 29 09:34:15 2003

OS Homo sapiens.
 PN WO2002186332-A2.
 XX 07-MAR-2002.
 PD
 XX 01-SEP-2001; 2001WO-EP10074.
 PF
 XX 01-SEP-2000; 2000DE-1043826.
 XX 05-SEP-2000; 2000DE-1044543.
 PR
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX WPI; 2002-371829/40.
 PI
 XX Determining the degree of cytosine methylation in genomic DNA, useful
 DR for diagnosis and prognosis, comprises selective hybridization of
 XX amplicons from chemically treated DNA -
 PT
 XX Claim 12; 56pp + Sequence Listing; 56pp; German.
 PS
 XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 SQ Sequence 1079 BP; 400 A; 387 C; 124 G; 168 T; 0 other;
 Query Match 3.9%; Score 99.2; DB 24; Length 1079;
 Best Local Similarity 71.2%; Pred. No. 1.9e-15;
 Matches 131; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
 QY 1 CGGGGCTAGCCCGGAGACCGCGCCGCTGGGGCGCTTACGGCGCTCTCGGACGG 60
 DB 724 CGAAGCTTAACCCGAAACCGGACCGGACCTTAACAGCCCTTCAGCGGTCTCGAAGCA 783
 QY 61 ATAATCGGTGAGCAGCAGCAGCGCGGAGCTCGGTGATCTGCCAGCAGCGCAGG 120
 DB 784 ATATACGATTAACAAACACACCGCGGACAACTCGACTTAATCTACGCAACAGCAAA 843
 QY 121 GATTGCGTGGCCCGGGAGCGCCGCGGAGCGGTGGGATCCTCAGCGCGCGCGGT 180
 DB 844 AATTACGTACCGCGGAAACCCGAAACACGACTAAATCTTCAACGAGCAGCGGATA 903
 QY 181 TGTC 184
 DB 904 AATC 907
 RESULT 9
 ID ABA46549
 AC ABA46549 standard; DNA; 413 BP.
 XX
 XX ABA46549;
 XX

DT 01-FEB-2002 (first entry)
 XX Human breast cell single exon nucleic acid probe #5244.
 DE Human; microarray; single exon probe; gene expression; breast;
 KW disease; cancer; ss.
 KW Homo sapiens.
 OS
 XX WO200157271-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00662.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-496933/54.
 DR
 XX New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -
 PT
 XX Claim 4; SEQ ID NO 5244; 327pp + sequence listing; English.
 PS
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a single exon nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 413 BP; 73 A; 131 C; 131 G; 78 T; 0 other;
 Query Match 3.3%; Score 85; DB 22; Length 413;
 Best Local Similarity 64.7%; Pred. No. 5.8e-12;
 Matches 143; Conservative 0; Mismatches 75; Indels 3; Gaps 1;
 QY 409 CAAATGCCAGCAAGACCCCAATTTTACCTGAAGAGCAGCAATAACAGAAAGAA 468
 DB 101 CACCATGTCACCAAGTGCCCATCTATCTGAAGCGTGGCAGTCGCAAG---GGCAAGAA 157
 QY 469 ATTTAACTGAGGACATCTCTCTCTGATATATCATGTCCTCCCGCTTGGAGACTTTCG 528
 DB 158 GGAGAAGCTTCGGGACCTGCTCTCTCGGACATATATAGCCACCGCTGGGGACTTCCG 217
 QY 529 CCACACCATCCCATTTGGCAAGAGGGCCAGCAGCATCTTTTGGAGATATTTCTTCT 588
 DB 218 CCACACCATTCATTTGGCGAGTGGCGGCGGAGTGCATGTTTGGGACATCTCTCTCT 277
 QY 589 TCAGGGAACCTACGAGCTTTTACCTGGAACACCGAGGAGAAAG 629

Db 278 GCAGGCAAGTTCACCTCTCTCGCGGACCATGGTGGAGG 318

RESULT 10
ABA64411

ID ABA64411 standard; DNA; 413 BP.

XX AC ABA64411;

DT 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #12716.

DE Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

KW Homo sapiens.

OS WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver

XX Claim 4; SEQ ID NO 12716; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

XX measuring human gene expression in a sample derived from human

XX liver. The single exon nucleic acid probes may be used for predicting,

XX measuring and displaying gene expression in samples derived from human

XX fetal liver. The present sequence is a single exon nucleic acid

XX probe of the invention.

XX Note: The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 413 BP; 73 A; 131 C; 131 G; 78 T; 0 other;

Query Match 3.3%; Score 85; DB 22; Length 413;

Best Local Similarity 64.7%; Pred. No. 5.8e-12;

Matches 143; Conservative 0; Mismatches 75; Indels 3; Gaps 1;

QY 409 CAAATGCCAGCAAGACCCCAATTTACCTGAAGCAGCCCAATACAAAGAAAGAA 468

DB 101 CACCATTGCCAAGGTGCCATCTATCTGAAGCGTGGCAGTCGCAAG---GGCAAGAA 157

QY 469 ATTTAACTGAGGACATCTCTCTGATATGATCAGTCCCGCTTGGAGACTTTCG 528

DB 158 GGAGAGCTTGGGACCTCTCTCTCGGACATGATCAGCCACCGCTGGGGGACTTCCG 217

QY 529 CCACCATCCACATTTGGCAAGAGGCGCCAGCAGCATGCTTTGGAGATATTTCTTCT 588

DB 218 CCACCATTCATATTGGCAGTGGCGGCGCAGTGCATGTTTGGGACATCTCTTCT 277

QY 589 TCAGGGAACACAGCTTTTACCTGGAACACAGGAGAAAG 629

DB 278 GCAGGCAAGTTCACCTCTCTCGCGGACCATGGTGGAGG 318

Db 278 GCAGGCAAGTTCACCTCTCTCGCGGACCATGGTGGAGG 318

RESULT 11
AAK12870

XX ID AAK12870 standard; DNA; 413 BP.

XX AC AAK12870;

XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe SEQ ID NO: 12861.

DE Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX epilepsy; cancer; ss.

XX OS Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains

XX Example 4; SEQ ID NO: 12861; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX brain. They can be used to measure gene expression in brain cell samples,

XX which may enable the diagnosis and improved treatment of nervous system

XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX epilepsy and cancers. The present sequence is one of the probes of the

XX SQ Sequence 413 BP; 73 A; 131 C; 131 G; 78 T; 0 other;

Query Match 3.3%; Score 85; DB 22; Length 413;

Best Local Similarity 64.7%; Pred. No. 5.8e-12;

Matches 143; Conservative 0; Mismatches 75; Indels 3; Gaps 1;

QY 409 CAAATGCCAGCAAGACCCCAATTTACCTGAAGCAGCCCAATACAAAGAAAGAA 468

DB 101 CACCATTGCCAAGGTGCCATCTATCTGAAGCGTGGCAGTCGCAAG---GGCAAGAA 157

QY 469 ATTTAACTGAGGACATCTCTCTGATATGATCAGTCCCGCTTGGAGACTTTCG 528

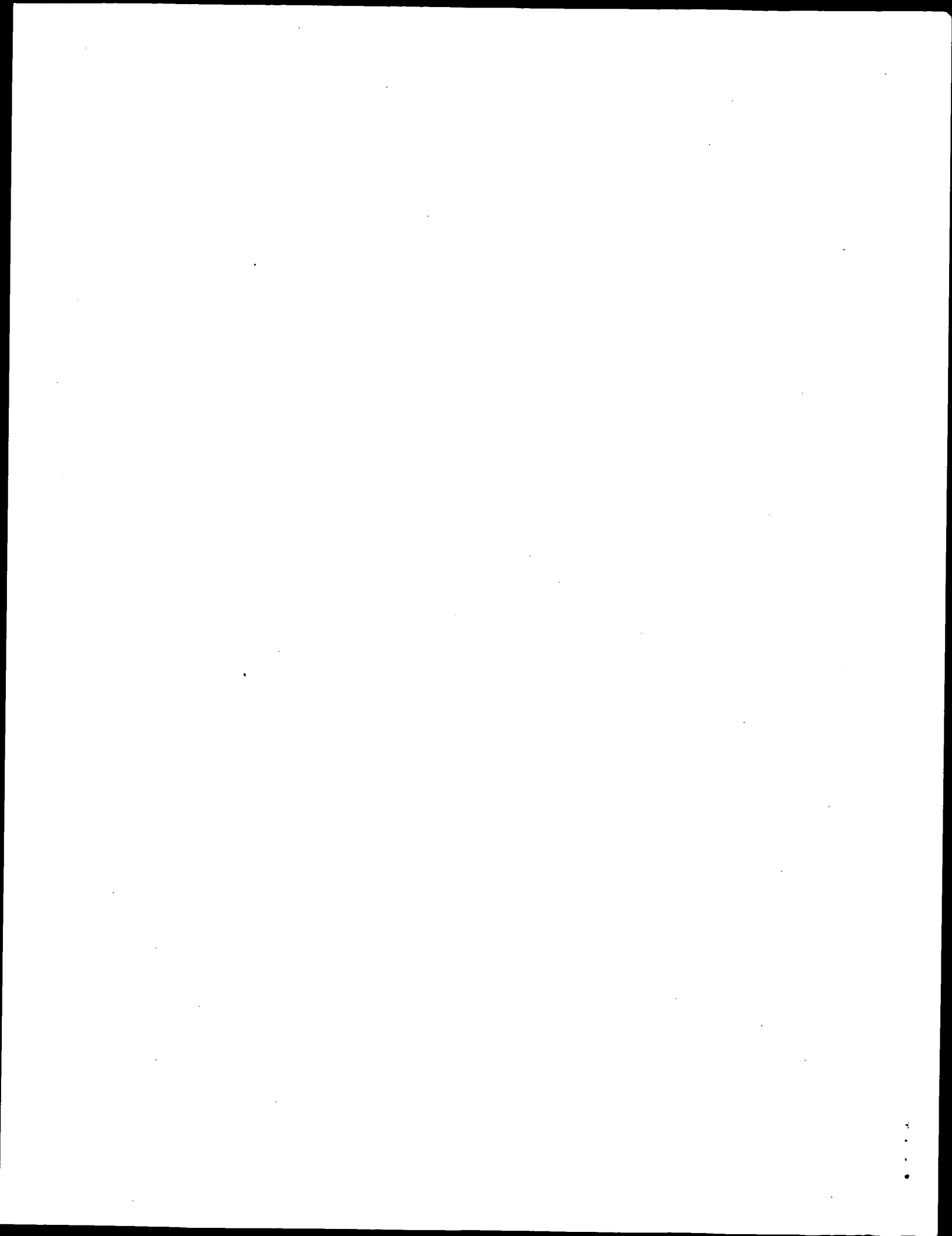
DB 158 GGAGAGCTTGGGACCTCTCTCTCGGACATGATCAGCCACCGCTGGGGGACTTCCG 217

QY 529 CCACCATCCACATTTGGCAAGAGGCGCCAGCAGCATGCTTTGGAGATATTTCTTCT 588

DB 218 CCACCATTCATATTGGCAGTGGCGGCGCAGTGCATGTTTGGGACATCTCTTCT 277

QY 589 TCAGGGAACACAGCTTTTACCTGGAACACAGGAGAAAG 629

DB 278 GCAGGCAAGTTCACCTCTCTCGCGGACCATGGTGGAGG 318



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 14:39:25 ; Search time 1425.45 Seconds
(without alignments)
10929.928 Million cell updates/sec

Title: US-09-806-276A-3
Perfect score: 962
Sequence: 1 tcgagccgattcgctcgag.....gaatctttgcaaaaaaaa 962

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_estl.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_man.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rtd.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	803.2	83.5	962	9	AL554570
2	787.6	81.9	897	14	BQ707991
3	787.6	81.9	923	14	BQ709678
4	783.2	81.4	960	14	BQ711479
5	782.8	81.4	888	14	BQ710002
6	771	80.1	972	14	BQ706417

7	768	79.8	926	14	BQ712054
8	765.8	79.6	924	12	BG745886
9	764.2	79.4	931	14	BQ707407
10	758.2	78.8	911	14	BQ899173
11	754.4	78.4	974	14	BQ707037
12	747.2	77.7	950	14	BQ709764
13	743	77.2	888	14	BQ706903
14	738.4	76.8	913	14	BQ711070
15	732	76.1	869	12	BG754043
16	728.4	75.7	938	14	BQ709175
17	725.6	75.4	841	12	BG758779
18	723.2	75.2	914	14	BQ707157
19	720.6	74.9	923	14	BQ711051
20	717.8	74.6	928	14	BQ709375
21	717.6	74.6	881	14	BQ709375
22	714	74.2	956	14	BQ706941
23	713.4	74.2	1007	14	BM914510
24	708.4	73.6	840	12	BG685798
25	708	73.6	849	13	BM007345
26	707	73.5	917	14	BQ711815
27	706.8	73.5	852	12	BG756918
28	706.4	73.4	975	14	BQ707855
29	704	73.2	851	12	BG754860
30	703.6	73.1	988	12	BG756209
31	703.4	73.1	893	12	BG397512
32	699.6	72.7	1075	12	BG758924
33	699	72.7	1010	14	BQ711412
34	697.8	72.5	924	12	BG745784
35	697	72.5	938	14	BQ709467
36	696.2	72.4	836	13	B1764715
37	694.8	72.2	858	13	B1764309
38	694.2	72.2	933	14	BQ711272
39	691.6	71.9	923	14	BQ709196
40	690.4	71.8	852	12	BG745626
41	690.2	71.7	894	12	BG746247
42	690	71.7	810	9	AL545157
43	689.2	71.6	914	13	BM007620
44	688.6	71.6	862	12	BG758749
45	687.6	71.5	844	12	BG685433

ALIGNMENTS

RESULT 1
AL554570 962 bp mRNA linear EST 16-FEB-2001
LOCUS LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI083YJ08 5
DEFINITION prime, mRNA sequence.
ACCESSION AL554570
VERSION AL554570.1 GI:12895480
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 962)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
FEATURES
source
1..962
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI083YJ08"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center drive Rockville, Maryland 20859, USA. Fax: (1) 301 610 8371 Email: liang@lifetech.com URL: <http://length.lifetech.com>

```

BASE COUNT      251 a      281 c      219 g      204 t      7 others
COUNT

```

Query Match	83.5%;	Score 803.2;	DB 9;	Length 962;
Best Local Similarity	93.8%;	Pred. No. 3.4e-192;		
Matches 875;	Conservative 6;	Mismatches 44;	Indels 8;	Gaps 4;
QY	36	TAGGACCCAGAGGGNACCATGGAGCCCGCAGCTCAGCTTCCTTCCTTCCTGCTACTCTGG	95	
Db				
1	TAGGACCCAGAGGGNACCATGGAGCCCGCAGCTTCCTTCCTTCCTGCTACTCTGG	59		
QY	96	CTCCAGATACACCGGAGAAATTTGTGTGACACAGTCTCCAGCCACCCCTGCTTGTGCT	155	
Db				
60	CTCCAGATACACCGGAGAAATTTGTGACGAGT-TCCAGGCACCCCTGCTTGTGCT	118		
QY	156	CCAGGGNAGAGCCACCTCTCTCGAGGGCCAGTCAGAGTGTA---GCAGCTACTTA	212	
Db				
119	CCCGGGACAGAGCCACCTCTCTCGAGGGCCAGTTAGAGTGTGATGATGATCAAC	272		
QY	213	GCCTGGTACCAAGAAACCTGGCCAGGCTCCAGGCTCTCATCTATGATGATCAAC	272	
Db				
179	GCCTGGTACCAGCAGAGCCTGGCCAGGCTCCAGAGCTCTCATCTATGATGATCAAC	238		
QY	273	AGGCCCACTGGCATCCACCCAGGTTCACTGGCAGTGGGTCTGGGACAGACTTCACTCTC	332	
Db				
239	AGGCCCACTGGCATCCAGACAGGTTCACTGGCAGTGGGTCTGGGACAGACTTCACTCTC	298		
QY	333	ACCATCAGCAGACTGGAGCCGAGAGTGTGGCAGTTTATTACTGTCAAGAAATTTT---	389	
Db				
299	ACCATCAGCAGACTGGAGCCTGAGGATTTTGCAGTTTACTGTCAAGAGTATGATGAC	358		
QY	390	ACTACTCCGTACACTTTTGGCCAGGGGACAGGCTGGAGTCAAGCAACTGTGGCTGCA	449	
Db				
359	TCACCTCGACTTACTTTCCGCGGGGACCAAGGTGGAGATCAAGCAACTGTGGCTGCA	418		
QY	450	CCATCTGTCTTCACTTCCCGCATCTGATGACAGTTTGAATCTGGAATCTGCCTGTGT	509	
Db				
419	CCATCTGTCTTCACTTCCCGCATCTGATGACAGTTTGAATCTGGAATCTGCCTGTGT	478		
QY	510	GTGTGCTGTGAATAACTTCTATCCAGAGAGGCCAAAGTACAGTGAAGTGGATAC	569	
Db				
479	GTGTGCTTCTGAATAACTTCTATCCAGAGAGGCCAAAGTACAGTGAAGTGGATAC	538		
QY	570	GCCTCCAAATCGGGTAACTCCAGAGAGTGTACAGAGCAGGACAGCAAGCAGCACC	629	
Db				
539	GCCTCCAAATCGGGTAACTCCAGAGAGTGTACAGAGCAGGACAGCAAGCAGCACC	598		
QY	630	TACAGCCTCAGCAGCACCCTGAGCTGAGCAAGCAGACTACGAGAAACACAAAGTCTAC	689	
Db				
599	TACAGCCTCAGCAGCACCCTGAGCTGAGCAAGCAGACTACGAGAAACACAAAGTCTAC	658		
QY	690	GCTTGGGAAGTCACCCATCAGGGCTGAGCTGCCTGTCACAAAGAGCTTCAACAGGGGA	749	
Db				
659	GCTTGCVAAGTCACCCATCAGGGCTGAGCTGCCTGTCACAAAGAGCTTCAACAGGGGA	718		
QY	750	GAGTGTATAGAGGAGAAAGTCCCCCACCCTGCTCTCAGTTCCAGCCTGACCCCCCTCCAT	809	
Db				
719	GAGTTTTATAGAGGAGAAAGTCCCCCACCCTGCTCTCAGTTCCAGCCTGACCCCCCTCCAT	778		
QY	810	CCTTTGGCCTTGACCCCTTTTCCACAGGGGACCTTACCCCTATTGCGGTCTCCAGTCA	869	
Db				
779	CCTTTGGCCTTADACCTTTTTCACAGGGGACCTTACCCCTATTGCGGTCTCCAGTCA	838		
QY	870	TCCTTTCACTTACCCCCCTCTCTCTCTTTAATATGCTTAATGTTGGAGGAAAT	929	

839	TCCTTTACCTCACCCCCCTCCTCCTCTGGCTTTAATTATGCTAATGTTGGAGGAGAAT	898
Qy	930	GAATAAATAAAGTGAATCTTTGCAAAAAA 962
Db	899	GAATAAATAAAGTKAATCTTTGCAAAAAA 931
RESULT 2		
LOCUS	BQ0707991	897 bp mRNA
DEFINITION	AGENCOURT_8347418 NIH_MGC_113 Homo sapiens	linear EST 16-JUL-2002
	5', mRNA sequence.	cdna clone IMAGE:6279302
ACCESSION	BQ0707991	
VERSION	BQ0707991.1	GI:21846890
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 897)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue Procurement: Dr. Mark Watson cdna Library Preparation: Rubin Laboratory cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2468 row: k column: 15 High quality sequence stop: 689.	

FEATURES
SOURCE

```

1. .897
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6279302"
/clone_lib="NIH_MGC_l13"
/lab_host="DH10B (phase-resistant)"
/note="organ: spleen; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
209 a 276 c 218 t 194 t
BASE COUNT

```

BASE COUNT	209 a	276 c	218 g	194 t
ORIGIN				
Query Match	81.9%;			
Best Local Similarity	94.5%;			
Matches 849;	Score 787.6;	DB 14;	Length 897;	
Conservative	Pred. No. 2.9e-188;			
	Mismatches 0;	Indels 5;	Gaps 3;	

32	CAGTTAGAACCCAGAGAGGAACCATGGAAGCCCGCAGCTCAGCTTCTCTTCTCCTCGTACT	91
1	CAGTTAGAACCCAGAGAGGAACCATGGAAGCCCGCAGCTTCTCTTCTCCTCGTACT	60
92	CTGGCTCCCGATACCAACCGGAGAAATGTGTGACACAGCTCCAGCCACCCCTGCTTTT	151
61	CTGGCTCCCGATACCAACCGGAGAAATGTGTGACGAGCTCCAGCCACCCCTGCTTTT	120
152	GTCTCCAGGGGAAAGGCCACCTCTCCTCAGGGCCAGTCAGAGTGTAGAGC---	208
121	GTCTCCAGGGGAAAGGCCACCTCTCCTCAGGGCCAGTCAGAGTATTAGCAGCAGTTA	180
209	CTTAGCCTGTGTACCAACAGAAACCTCGCCAGGCTCCCGAGGCTCCTCATCTATGATGCATC	268
181	CATAGCCTGTGTATCAGGAGAAACCTGGCCAGGCTCCCGAGGCTCCTCATCTATGATGCATC	240
269	CAACAGGGCCACTGGCATCCACCCAGTTCTAGTGGCAGTGGGTCTGGGACAGACTTCAC	328

Db 541 TAAGCCCTCCATCGGTAACCTCCAGGAGGTGTACAGAGCAGGACGACAGCAG 600
 Qy 626 CACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAGCAGACTACGAGAAACACAAAGT 685
 Db 601 CACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAGCAGACTACGAGAAACACAAAGT 660
 Qy 686 CTAGCCTGGGAGTACCATCAGGCGCTGAGCTGCCGCTCACAAAGAGCTTCAACAG 745
 Db 661 CTAGCCTGGGAGTACCATCAGGCGCTGAGCTGCCGCTCACAAAGAGCTTCAACAG 720
 Qy 746 GGGAGAGTGTAGAGGAGAGTCCGCCACCTGCTCTCTCAGTTCAGCCTGACCCCTC 805
 Db 721 GGGAGAGTGTAGAGGAGAGTCCGCCACCTGCTCTCTCAGTTCAGCCTGACCCCTC 780
 Qy 806 CCATCCTTTGGCTCTGACCTTTTCCACAGGGGACCTACCCCTATTGGGGTCTCCAG 865
 Db 781 CCATCCTTTGGCTCTGACCTTTTCCACAGGGGACCTACCCCTATTGGGGTCTCCAG 839
 Qy 866 CTCATCTTTACCTCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 925
 Db 840 CTCATCTTTACCTCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 898
 Qy 926 GAATGAATAAATAAA 940
 Db 899 GGATGAATAAATAAA 913

RESULT 8
 BG745886/c
 LOCUS
 DEFINITION 602724056r1 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:4850557 3',
 mRNA sequence.
 ACCESSION BG745886
 VERSION BG745886.1 GI:14056539
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 924)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Dr. Mark Watson
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM1691 row: p column: 14
 High quality sequence start: 2
 High quality sequence stop: 918.
 Location/Qualifiers
 1..924
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4850557"
 /lab_host="PH10B (phage-resistant)"
 /note="Organ: spleen; Vector: pOTB7; Site: 1: xhoI; Site: 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

FEATURES
source

BASE COUNT 202 a 216 c 283 g 223 t
 ORIGIN
 RESULT 9
 BG707407
 LOCUS
 DEFINITION AGENCOURT_8495191 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6302024

Query Match 79.6%; Score 765.8; DB 12; Length 924;
 Best Local Similarity 93.8%; Pred. No. 9.4e-183;
 Matches 864; Conservative 0; Mismatches 47; Indels 10; Gaps 6;
 Qy 29 GCTCAGTTAGACCCAGAGGAAACCATGTAAGCCACCCAGCTCAGCTTCTCTCTCTCTCTCT 88
 Db 924 GCTCAGTTAGACCCAGAGGAAACCATGTAAGCCACCCAGCTCAGCTTCTCTCTCTCTCTCT 865
 Qy 89 ACTCTGGCTCCG-AGATACCACCGGAGAAATTTGTTGACACAGTCTCCAGCCACCCCTGT 147
 Db 864 ACTCTGGCTCCG-AGATACCACCGGAGAAATTTGTTGACACAGTCTCCAGCCACCCCTGT 805
 Qy 148 CTTTGTCTCCAGGGG-AAAGAGCCACCCCT--CTCCTGACGGGCCAGTACAGAGTGT--A 201
 Db 804 CTCTGTCTCCAGGGGTAAGAGACCCCTGCTCCCTGACAGGCCAGTACAGAGTGTAGCA 745
 Qy 202 GCAGCTACTTAGCTGGTACCACACAAACCTTGGCCAGGCTCCAGGCTCTCTCTCTCTCTCT 261
 Db 744 GCAGCTACTTAGCTGGTACCACACAAACCTTGGCCAGGCTCCAGGCTCTCTCTCTCTCTCT 685
 Qy 262 ATGCATCCAAACAGGGCCACTGGCATCCACCCAGGTTTCAGTGGCAGTGGGTCTGGGACAG 321
 Db 684 GTGCATCCAGCAGGCTCCACTGGCATCCACAGACAGGTTTCAGTGGAAATGGTTCAGGACAG 625
 Qy 322 ACTTCACTCTCACCATCAGCAGACTGGAGCCGCAAGATGTGGCACTTATTACTGTCTCAGC 381
 Db 624 ACTTCACTCTCACCATCAGCAGACTGGAGCCGCTCAGATTTTGCAGTGTATTACTGTCTCAGC 565
 Qy 382 AATATTTTACTACTCCGTACACTTTTGGCCAGGAGCAGGCTGGAGATCAAAACGAACGTG 441
 Db 564 AGTATGAAACTCAGTATTACCTTTTCGGCCCTGGGACCAAGTGGATATCAAAACGAACGTG 505
 Qy 442 TGGCTGCACCATCTGTCTTTCATCTTCCCGCCATCTGATGAGCAGTGTGAATCTGGAACGTG 501
 Db 504 TGGCTGCACCATCTGTCTTTCATCTTCCCGCCATCTGATGAGCAGTGTGAATCTGGAACGTG 445
 Qy 502 CTTCTGTGTGTGCTGTCTGAATAACTTCTATCCAGAGAGGCCAAGTACAGTGGAAAG 561
 Db 444 CTTCTGTGTGTGCTGTCTGAATAACTTCTATCCAGAGAGGCCAAGTACAGTGGAAAG 385
 Qy 562 TGGATAACGCCCTCCCAATCGGTAACCTCCAGGAGAGTGTACAGAGCAGGACAGCAAGG 621
 Db 384 TGGATAACGCCCTCCCAATCGGTAACCTCCAGGAGAGTGTACAGAGCAGGACAGCAAGG 325
 Qy 622 ACAGCACCTACAGCCTCAGCAGCAACCTGACGCTGAGCAAGAGCAGTACGAGAAACACA 681
 Db 324 ACAGCACCTACAGCCTCAGCAGCAACCTGACGCTGAGCAAGAGCAGTACGAGAAACACA 265
 Qy 682 AAGTCTACGCTGGGAAGTCAACCATCAGGCGCTGAGCTGCCGCTCACAAAGAGCTTCA 741
 Db 264 AAGTCTACGCTGGGAAGTCAACCATCAGGCGCTGAGCTGCCGCTCACAAAGAGCTTCA 205
 Qy 742 ACAGGGAGAGTGTAGAGGAGAGTGGCCGCCACCTGCTCTCTCAGTTCACAGCCTGACCC 801
 Db 204 ACAGGGAGAGTGTAGAGGAGAGTGGCCGCCACCTGCTCTCTCAGTTCACAGCCTGACCC 145
 Qy 802 CCTCCCATCTTTGGCCTCTGACCTTTTCCACAGGGGACCTTACCCCTATTGGGGTCTCT 861
 Db 144 CCTCCCATCTTTGGCCTCTGACCTTTTCCACAGGGGACCTTACCCCTATTGGGGTCTCT 85
 Qy 862 CCAGCTCATCTTTTCACTCACCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 919
 Db 84 CCAGCTCATCTTTTCACTCACCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26
 Qy 920 GGAGGAGAAATGAATAAATAAA 940
 Db 25 GGAGGAGAAATGAATAAATAA 5

RESULT 9
 BG707407
 LOCUS
 DEFINITION AGENCOURT_8495191 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6302024

931 bp mRNA linear EST 16-JUL-2002
 B0707407
 AGENCOURT_8495191 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6302024

```

5', mRNA sequence.
BQ707407
VERSION BQ707407.1 GI:21846306
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2518 row: n column: 09
High quality sequence start: 17
High quality sequence stop: 656.
FEATURES
    source
        1..931
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:6302024"
            /lab_host="NIH_MGC_113"
            /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G). Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: this is a
            NIH_MGC Library."
BASE COUNT 219 a 283 c 223 g 206 t
ORIGIN
Query Match 79.4%; Score 764.2; DB 14; Length 931;
Best Local Similarity 92.1%; Pred No. 2.4e-182;
Matches 841; Conservative 0; Mismatches 63; Indels 9; Gaps 3;
QY 42 CCAGAGGGAAACCATGGAAGCCAGCTCAGCTCTCTTCCCTGCTACTCTGGCTCCCA 101
DB 10 CCAGAGGGAAACCATGGAAGCCAGCTCTCTTCCCTGCTACTCTGGCTCCCA 69
QY 102 GATACACCGGAGAAATGTTGTGACACAGTCTCCAGCCAGCTCTTCTCTCCAGGG 161
DB 70 GATACACCGGAGAAATGCGTTGACGAGTCTCCAGCCAGCTCTTCTCTCCAGGG 129
QY 162 GAAGAGGCCACCTCTCTGAGGCCAGTCTGAGTGTAGCAG--CTACTTACGCTGG 218
DB 130 GAAGAGGCCACCTCTCTGAGGCCAGTCTGAGTGTAGCAG--CTACTTACGCTGG 189
QY 219 TACCACAGAAACCTGCGCCAGCTCTCCAGCTCTCTATGATGATGATGATGATGATG 278
DB 190 TACCACAGAAATCTGGCCAGCTCTCCAGCTCTCTATGATGATGATGATGATGATG 249
QY 279 ACTGGCATCCACCCAGCTCTGAGTGGAGTGGTCTGGAGAGACTTCACTCTCACCATC 338
DB 250 GCTGTCATCCACAGAGCTCTGAGTGGTGGTCTGGAGAGACTTCACTCTCACCATC 309
QY 339 AGCAGACTGGAGCCCGAAGATGTGGCACTTTATTACTGTGACGAATATTTACTACTCCG 398
DB 310 AGCAGACTGGAGCCCGAAGATGTGGCACTTTATTACTGTGACGAATATTTACTACTCCG 369
QY 399 TACACTTTTGGCCAGGAGCCAGCTGGAGATCAACAGCACTGTGGCTGCACCATCTGTC 458
DB 370 TTCACCTTTGGCCGAGCAAGATGTGGATATCAAGCAAGCACTGTGGCTGCACCATCTGTC 429

```

```

QY 459 TTCATCTTCCGCCCATCTGATGACGAGTTGAATCTGGAAGTGGCTGTGTGTGCTGCTG 518
DB 430 TTCATCTTCCGCCCATCTGATGACGAGTTGAATCTGGAAGTGGCTGTGTGTGCTGCTG 489
QY 519 CTGAATAACTTCTATCCAGAGAGGCCAAAGTACAGTGGAGTGGGATAAGCCCTCCAA 578
DB 490 CTGAATAACTTCTATCCAGAGAGGCCAAAGTACAGTGGAGTGGGATAAGCCCTCCAA 549
QY 579 TCGGTAATACTCCAGGAGAGTGTACAGAGCAGGACAGCAAGAGACACCTACAGCTC 638
DB 550 TCGGTAATACTCCAGGAGAGTGTACAGAGCAGGACAGCAAGAGACACCTACAGCTC 609
QY 639 AGCAGCACCTGAGCGCTGAGCAAAAGCAGACTACAGAGAAACACAAGTCTAGCCCTGGAA 698
DB 610 AGCAGCACCTGAGCGCTGAGCAAAAGCAGACTACAGAGAAACACAAGTCTAGCCCTGGAA 669
QY 699 GTCACCCATCAGGGCCTGAGCTCGCCGCTCACAAGAGCTTCAACAGGGGAGAGTGTAG 758
DB 670 GTCACCCATCAGGGCCTGAGCTCGCCGCTCACAAGAGCTTCAACAGGGGAGAGTGTAA 729
QY 759 AGGAGAGAGTCCCGCCACCTGCTCTCAGTTCCAGCTGACCCCTCCCATCTTTGGCC 818
DB 730 AGGAGAGAGTCCCGCCACCTGCTCTCAGTTCCAGCTGACCCCTCCCATCTTTGGCC 789
QY 819 TCTGACCTTTTCCACAGGGGACCTA-CCCTATTTCGGTCTCTCCAGCTCATCTTTCAC 877
DB 790 TCTGACCTTTTCCACAGGGGACCTA-CCCTATTTCGGTCTCTCCAGCTCATCTTTCAC 849
QY 878 CTCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 932
DB 850 ACCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 909
QY 933 TAAATAAGTGAA 945
DB 910 TAAATAAGGGGA 922
RESULT 10
BQ899173
LOCUS BQ899173
DEFINITION BQ899173 911 bp mRNA linear EST 16-AUG-2002
5', mRNA sequence.
ACCESSION BQ899173
VERSION BQ899173.1 GI:22291187
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2519 row: i column: 15
High quality sequence start: 4
High quality sequence stop: 679.
FEATURES
    source
        1..911
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:6302294"
            /lab_host="NIH_MGC_113"
            /lab_host="DH10B (phage-resistant)"

```

/note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGCAG(C). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a XhoI-MCC library."

[illegible]

DB	841	TACCTTACCCCTCCTCCCTCTTGCGCTTAATTTATGCCAAAGTTGGAGGAAAATG	900
QY	931	AATAATAAAG 941	
DB	901	GATAATAAAG 911	
RESULT 11			
BO707037			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
COMMENT			

FEATURES

```

1. .974
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_113"
/lab_host="PHI08 (phage-resistant)"
/note="Organ: spleen; vector: pOM7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
244 a      284 c      237 g      207 t      2 others
BASE COUNT
ORIGIN
Query Match          78.4%; Score 754.4; DB 14; Length 974;
Best Local Similarity 93.6%; Pred. No. 7.2e-180;
Matches              798; Conservative           0; Mismatches 52; Indels    3; Gaps   1;

Qy      48  GGAAACCATGTGAAGCCCGCAGCTCAGCTTCCTCTCTCTGCTACTTGCTGGCTCCGATACC 107
       |||
Db      21  GGAACATGGAACCCCGCAGCTTCTTCTCTGCTACTTGCTGGCTCCGATACC 80
       |||

Qy     108  ACCGGAGAAATTGTGTGACAGCTCTCCAGCACCTGCTTTGCTTCAGGGGAAAAGA 167
       |||
Db     81  ACCGGAGAAATTGTGTTGACGCGFTCCAGGCACTGCTTTGCTTCAGGGGAAAAGA 140
       |||

Qy     168  GCCACCTCTCTCGAGGGCCAGTCAGAGTGT---AGCAGCTACTTAGCTGGTACCAG 224
       |||
Db    141  GCCACCTCTCTCGAGGGCCAGTCAGAGTGTTAGCAGCAGCTACTTAGCTGGTACCAG 200
       |||

Qy     225  CAGAAACCTGGCCAGGCTCCCAGGCTTCCTCATCTATGATCATCCAACAGGGCCACTGGC 284
       |||
Db    201  CAATAACCTGGCCAGGCTCCCAGGCTTCCTCATCTATGAGCATCCACAGGGCCACTGGC 260
       |||

Qy     285  ATCCACCCAGGTTTCAGTGGCAGTGGTCTGGGACAGAGCTTCACTCTCACCATCAGCAGA 344

```

db	261	ATCCAGACAGGTTTCAGTGGCGAGTGGCTGGGACAGACTTCACTCTCACCGTCAGACA	320
ay	345	CTGGAGCCGGAAGATGTGGCACTTTATTACTGTGCAGCAATATTTTACTACTCGGTACACT	404
db	321	CTGGAGCCTGAAGATTTTGCAGTGTATTACTGTGCAGCATGTGGTAGCTCAACGAATCACT	380
ay	405	TTTGGCCAGGGGACCGAGCTGGAGATCAAAAGAACTGTGGCTGCACCATCTGCTTCATC	464
db	381	TTGCGCCCTGGGACCAAAAGTGGAGATCAAAAGAACTGTGGCTGCACCATCTGCTTCATC	440
ay	465	TTCCGGCCATCTGATGAGCAGTTGAAATCTGGAACCTGCCCTCTGTTGTGTGCTGTGTAAT	524
db	441	TTCCGGCCATCTGATGAGCAGTTGAAATCTGGAACCTGCCCTCTGTTGTGTGCTGTGTAAT	500
ay	525	AACCTTCATCTCCAGAGAGGCCAAAGTACAGTGGAAAGTGGATTAACGCCCTCCAATCGGGT	584
db	501	AACCTTCATCTCCAGAGAGGCCAAAGTACAGTGGAAAGTGGATTAACGCCCTCCAATCGGGT	560
ay	585	AACCTCCAGGAGAGTGTACAGACAGACAGACAGCAAGACAGCACTACAGCCTCAGCAGC	644
b	561	AACCTCCAGGAGAGTGTACAGACAGGAGACAGCAAGACAGCACTACAGCCTCAGCAGC	620
y	645	ACCCCTGAGCTGAGCAAAAGCAGACTACGAGAAACACAAAGTCTAGCCCTGGGAAGTCAAC	704
b	621	ACCCCTGAGCTGAGCAAAAGCAGACTACGAGAAACACAAAGTCTAGCCCTGGGAAGTCAAC	680
y	705	CATCAGGGCTGAGTCTGCCCGTCAAAAGAGTTCACAGGGGAGAGTGTATAGAGGGAG	764
b	681	CATCAGGGCTGAGTCTGCCCGTCAAAAGAGTTCACAGGGGAGAGTGTATAGAGGGAG	740
y	765	AAGTGGCCCCACCTGCTCCTCAGTTCAGCTCAGCCCTGACCCCTCCCATCTTTGGGCTCTGAC	824
b	741	AAGGGCCCCACCTGCTCCTCAGTTCAGCTCAGCCCTGACCCCTCCCATCTTTGGGCTCTGAC	800
y	825	CTTTTTCACAGGGGACCTACCCCTATTGCGGTCCTCCAGCTCATCTTTTCACTCACC	884
b	801	CTTTTTCACAGGGGACCTACCCCTATTGCGGTCCTCCAGCTCATCTTTTCACTCACC	860
y	885	CGCTCCTCCTCCT	897
b	861	CCCCCTCCTCCT	873

RESULT 12					
B0709764					
LOCUS	B0709764	950 bp	mRNA	linear	EST 16-JUL-2002
DEFINITION	AGENCOURT_8352455	NIH_MGC_113	Homo sapiens	cdna clone	IMAGE:6277649
				5'	mRNA sequence.

BO709764.1 GI:21848663
 EST.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 950)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email:

Tissue Procurement: Dr. Mark Watson
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Cloning Distribution by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM2464 row: f column: 18
 High quality sequence stop: 641.
 Location/Qualifiers
 1..950
 FEATURES
 source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6277649"
/clone_lib="NIH_MGC_113"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: spleen; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
238 a 275 c 228 g 209 t

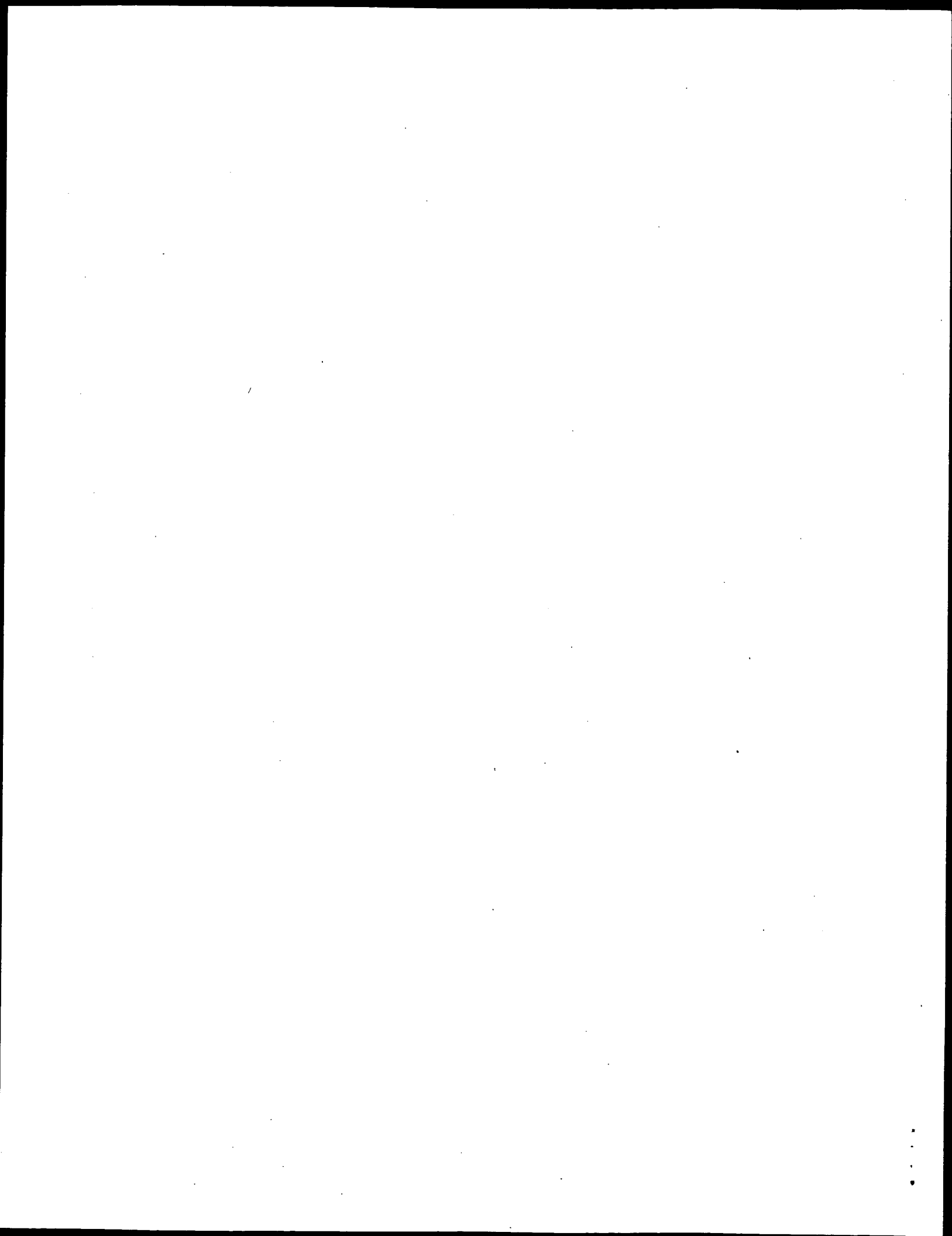
```

[illegible]

•


```
|||||
Db 121 CTCCTGCGAGGCGAGTCTAGAGTGTAGCAGCAGCTACTTACCTGGTACCGACGAGAAACC 180
QY 233 TGGCCAGGCTCCCGAGGCTCCCTCATCTATGATGATCATCAACACAGGCGCACTGGCATCCCAACC 292
Db 181 TGGCCAGGCTCCCGAGGCTCCCTCATCTATGATGATCATCAACACAGGCGCACTGGCATCCCAACC 240
QY 293 CAGGTTCAAGTGGCAGTGGGCTCTGGGACAGAGTTTCACTCTCACCATCAGCAGACTGGAGCC 352
Db 241 CAGGTTCAAGTGGCAGTGGGCTCTGGGACAGAGTTTCACTCTCACCATCAGCAGACTGGAGCC 300
QY 353 CGAAGATGTGGCAGCTTTTATCTGTCAGCAATATTTTACTACTCCGTACACTTTTGGCCA 412
Db 301 TGAAGATTTTGCACGTGATTTACTGTCAGCAGTATGGTAGCTCACCGTACACTTTTGGCCA 360
QY 413 GGGGACCGAGGCTGGAGATCAACAGAACTGTGGCTGCACCATCTGTCTTCACTCTTCCCGGCC 472
Db 361 -GGGACCAAGCTGGAGATCAACAGAACTGTGGCTGCACCATCTGTCTTCACTCTTCCCGGCC 419
QY 473 ATCTGATGAGCAGTTGAATCTGGAACTGCCCTCTGTTGCTGCTGCTGAATAACTTCTA 532
Db 420 ATCTGATGAGCAGTTGAATCTGGAACTGCCCTCTGTTGCTGCTGCTGAATAACTTCTA 479
QY 533 TCCGAGAGAGGCGCAAGTACAGTGAAGGTGGATAACGCCCTTCCAATCGGTTAACTCCCA 592
Db 480 TCCGAGAGAGGCGCAAGTACAGTGAAGGTGGATAACGCCCTTCCAATCGGTTAACTCCCA 539
QY 593 GGAGAGTGTCAAGAGCAGGAGCAGCAAGGACAGCAGCTTACAGCCTCAGCAGCAGCCTGAC 652
Db 540 GGAGAGTGTCAAGAGCAGGAGCAGCAAGGACAGCAGCTTACAGCCTCAGCAGCAGCCTGAC 599
QY 653 GCTGAGCAAGCAGAGCTAGGAGAAACACAAAGCTTACGCTGCGAAGTCAACCATCAGGG 712
Db 600 GCTGAGCAAGCAGAGCTAGGAGAAACACAAAGCTTACGCTGCGAAGTCAACCATCAGGG 659
QY 713 CCTGAGCTCGCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTAGAGGG----AGAAGT 768
Db 660 CCTGAGCTCGCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTAGAGGGACCAAGGTGC 719
QY 769 GCGCCGACCTGCTCTCAGTT--CCAGCCTGACCCGCTCCCATCTTGGCCTCTGACCC 826
Db 720 CCCCCAACTTGCTCTCAGTTTCCAGCCTTGACCCCTTCCAAATCCTTTGGCCTCTGACCC 779
QY 827 TTTTCCACAGGGGACCTACCCCTATTGGGTCCTCCAGCTCATCTTTACCTCACCCCC 886
Db 780 TTTTCCACAGGGGACCTAGCCCTATTGGGGCTCCAGCTCATCTTTACCTCACCCCT 839
QY 887 CTCCTCCTCTTGGCTTTTAAATGCTAATGTTG 920
Db 840 CCTCTCC---TTGGTTTAAATGCTAATGTTG 869
```

Search completed: July 25, 2003, 19:53:44
Job time : 1434.45 secs



Tue Jul 29 09:34:12 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 15:39:40 ; Search time 212.803 Seconds
(without alignments)
9326.039 Million cell updates/sec

Title: US-09-806-276A-3
Perfect score: 962
Sequence: 1 tctgagcgcattgctgcag.....gaattttgcaaaaaaaa 962

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues
Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/ECT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/ECTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	850.4	88.4	1184	15	US-10-158-646-60	Sequence 60, Appl
2	845.4	87.9	970	11	US-09-859-053-37	Sequence 37, Appl
3	830.2	86.3	948	11	US-09-859-053-33	Sequence 33, Appl
4	816.4	84.9	1202	15	US-10-158-646-57	Sequence 57, Appl
5	808.4	84.0	913	11	US-09-822-830A-531	Sequence 531, Appl
6	792.4	82.4	1244	11	US-09-954-456-771	Sequence 771, Appl
7	767.2	79.8	1001	15	US-10-198-846-14013	Sequence 14013, A
8	746.8	77.6	974	11	US-09-859-053-29	Sequence 29, Appl
9	746.8	77.5	1775	15	US-10-158-646-64	Sequence 64, Appl
10	738.8	76.8	1033	10	US-09-799-514-2	Sequence 59, Appl
11	726.4	75.5	1230	15	US-10-158-646-59	Sequence 59, Appl
12	724.6	75.3	990	10	US-09-800-729-79	Sequence 79, Appl
13	716.8	74.5	1458	15	US-10-158-646-66	Sequence 66, Appl
14	713.8	74.2	961	15	US-10-198-846-13125	Sequence 13125, A
15	713.2	74.1	941	15	US-09-800-729-81	Sequence 81, Appl
16	710.4	73.8	928	15	US-10-221-945-5	Sequence 5, Appl

Sequence 4, Appl
Sequence 68, Appl
Sequence 13629, A
Sequence 7, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 64, Appl
Sequence 63, Appl
Sequence 63, Appl
Sequence 14, Appl
Sequence 13869, A
Sequence 6, Appl
Sequence 65, Appl
Sequence 10845, A
Sequence 10, Appl
Sequence 13859, A
Sequence 83, Appl
Sequence 81, Appl
Sequence 87, Appl
Sequence 89, Appl
Sequence 35, Appl
Sequence 45, Appl
Sequence 13823, A
Sequence 85, Appl
Sequence 49, Appl
Sequence 73, Appl
Sequence 99, Appl

ALIGNMENTS

RESULT 1
US-10-158-646-60
; Sequence 60, Application US/10158646
; Publication No. US20030073105A1
; GENERAL INFORMATION:

APPLICANT: Lasek, Amy K.W.
TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
FILE REFERENCE: PA-0030-1 US
CURRENT APPLICATION NUMBER: US/10/158,646
CURRENT FILING DATE: 2002-05-29
PRIOR FILING DATE: 2001-05-31
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PERL Program
SEQ ID NO 60
LENGTH: 1184
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030073105A1 1135037.21
FEATURE:
NAME/KEY: unsure
LOCATION: 44, 65
OTHER INFORMATION: a, t, c, g, or other
US-10-158-646-60

Query Match 88.4%; Score 850.4; DB 15; Length 1184;
Best Local Similarity 96.2%; Pred. No. 1.2e-238;
Matches 893; Conservative 31; Indels 4; Gaps 2;
QY 29 GTCAGTTAGGACCCAGAGGGGACCATGGAGGCCCGCCAGCTTCCTCTCTCCCTCCCTGCT 88
Db 257 GTCAGTTAGGACCCAGAGGGGACCATGGAGGCCCGCCAGCTTCCTCTCTCCCTCCCTGCT 316
QY 89 ACTCTGGCTCCCGAGATACACCGGAGAAATTTGTTGACAGATC-TCAGCCACCCTGT 147
Db 317 ACTCTGGCTCCCGAGATACACCGGAGAAATTTGTTGACAGATC-TCAGCCACCCTGT 376

D _b	741	TAGAGGAGAAGTGC	CCCCACCTGCTCCTCAGTTCCAGGCTGACC	CCCCCTCCCATCCTTCG	800
Q _y	816	GGCTGTAGCCCTTTTT	CCACAGGGGACCTACCCCTATTGGGGTCC	TCCAGCTCATCTTTC	875
D _b	801	GGCTGTAGCCCTTTTT	CCACAGGGGACCTACCCCTATTGGGGTCC	TCCAGCTCATCTTTC	860
Q _y	876	ACCTCACCCCGCTCCT	CCTCCTTGCGCTTAATTTATGCTAATGTTGGAGGAGAA	TGAATAA	935
D _b	861	ACCTCACCCCGCTCCT	CCTCCTTGCGCTTAATTTATGCTAATGTTGGAGGAGAA	TGAATAA	920
Q _y	936	ATAAAGTGAATCTTT	GCAAAAAAAA	962	
D _b	921	ATAAAGTGAATCTTT	GCAAAAAAAA	947	

```

RESULT 3
US-09-859-053-33
; Sequence 33, Application US/09859053
; Patent No. US20020102658A1
; GENERAL INFORMATION:
; APPLICANT: Tsuji, Takashi
; APPLICANT: Tzuka, Katsunari
; APPLICANT: Hori, No. US20020102658A1uaki
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A
; TRANSDUCTION MOLECULE AILIM AND
; TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE AILIM AND
; TITLE OF INVENTION: PHARMACEUTICAL USE THEREOF
; FILE REFERENCE: 06501-079001
; CURRENT APPLICATION NUMBER: US/09/859,053
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: JP 2001-99508
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: JP 2000-147116
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(27)
; NAME/KEY: CDS
; LOCATION: (28)...(735)
; NAME/KEY: 3'UTR
; LOCATION: (739)...(948)
; NAME/KEY: sig_peptide
; LOCATION: (28)...(87)
; NAME/KEY: misc_feature
; LOCATION: (1)...(948)
; OTHER INFORMATION: n = A,T,C or G
US-09-859-053-33

```

289	QY	289	CACCAGGTTTCAGTGGCAGTGGGTCTGGACAGACTTCACCTCTCACCATCAGCAGACTGG	348
266	Db	266	CAGACAGGTTTCAGTGGCAGTGGGTCTGGACAGACTTCACCTCTCACCATCAGCAGACTGG	325
349	QY	349	AGCCCGAAGATGTGGCAGCTTATTAAGTGTGACAGAAATATTTTA--CTACTCCGTACACACTT	405
326	Db	326	AGCCTGAAGATTTTCAGATGTATTACTGTTCAGCAGTTTGGTAGCTCACCTATGTGCGATT	385
406	QY	406	TTGGCCAGGGGACCAGGCTGGAGATCAACAGAACTGTGGCTGCACCATCTGCTTTCATCT	465
386	Db	386	TTGGCCAGGGGACCAAGCTGGAGATCAACAGAACTGTGGCTGCACCATCTGCTTTCATCT	445
466	QY	466	TCCCGGCATCTGATGAGCAGTTGAAATCTGGAATCGCTTGTGTGTGCGCTGCTGTGAATA	525
446	Db	446	TCCCGCATCTGATGAGCAGTTGAAATCTGGAATCGCTTGTGTGTGCGCTGCTGTGAATA	505
526	QY	526	ACTTCTATCCACAGAGAGGCCAAAGTACAGTGGAAAGTGGATACGCGCTCCCAATCGGGTA	585
506	Db	506	ACTTCTATCCACAGAGAGGCCAAAGTACAGTGGAAAGTGGATACGCGCTCCCAATCGGGTA	565
586	QY	586	ACTCCAGGAGAGTGTACAGCAGCAGGACGACGACGAGGACGACCTACAGCCTCAGCAGCA	645
566	Db	566	ACTCCAGGAGAGTGTACAGCAGCAGGACGACGACGAGGACGACCTACAGCCTCAGCAGCA	625
646	QY	646	CCCTGACGCTTGACAAAGCAGACTACGAGAAACACAAGTCTACGCCCTGGGAGTCAACC	705
626	Db	626	CCCTGACGCTTGACAAAGCAGACTACGAGAAACACAAGTCTACGCCCTGGGAGTCAACC	685
706	QY	706	ATCAGGCGCTTGAGCTCGCCGCTCACAAAGAGCTTCAACAGGGAGAGTGTAGAGGGAGA	765
686	Db	686	ATCAGGCGCTTGAGCTCGCCGCTCACAAAGAGCTTCAACAGGGAGAGTGTAGAGGGAGA	745
766	QY	766	AGTGCCTCCACCTGCTCCTCAGTTCACAGCCTGACCCCTCCCATTCCTTTGGCCTCTGACC	825
746	Db	746	ANTGCCCTCCACCTGCTCCTCAGTTCACAGCCTGACCCCTCCCATTCCTTTGGCCTCTGACC	805
826	QY	826	CTTTTTCACAGGGGACCTTACCCCTATTTCGGGTCTCCAGCTCATCTTTCACCTCACCCC	885
806	Db	806	CTTTTTCACAGGGGACCTTACCCCTATTTCGGGTCTCCAGCTCATCTTTCACCTCACCCC	865
886	QY	886	CCTCCTCCTCTTGGGCTTTAAATATTGCTAATGTGGAGGAGAAATGAATAAATAAGTGAA	945
866	Db	866	CCTCCTCCTCTTGGGCTTTAAATATTGCTAATGTGGAGGAGAAATGAATAAATAAGTGAA	925
946	QY	946	TCTTTGCA	953
926	Db	926	TCTTTGCA	933

```

RESULT 4
US-10-158-646-57
; Sequence 57, Application US/10158646
; Publication NO. US20030073105A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Sornasse, Thierry
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0030-1 US
; CURRENT APPLICATION NUMBER: US/10/158,646
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/295,239
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PERL Program
; SEQ ID NO 57
; LENGTH: 1202
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030073105A1 1135037.27
US-10-158-646-57

```

Query Match 84.9%; Score 816.4; DB 15; Length 1202;
Best Local Similarity 95.7%; Pred. No. 1.1e-228;
Matches 907; Conservative 0; Mismatches 26; Indels 15; Gaps 6;

QY 29 GCTCAGTTAGACCCAGGAGGAGCCAT-GGAAGCCCGAGCTCAGCTTCTCTCTCTCTG 87
Db 256 GCTCAGTTAGACCCAGGAGGAGCCATGGGAAACCCCGAGGAGCTTCTCTCTCTCTG 315
QY 88 TACTCTGCTCC-----CAGATACCACCGAGAGAAATGTTGTGACACAGTC-TCAGGCC 140
Db 316 TACTCTGCTCCAGTCGCGAGTTATACCGGAGAAATGTTGTGACGAGCTCTCCAGGC 375
QY 141 ACCCTGCTTTGCTCTCAGGGGAAAGACCCCTCTCTGAGGGCCAGTCAGAGTGT 200
Db 376 ACCCTGCTTTGCTCTCAGGGGAAAGAGCCACCTCTCTGAGGGCCAGTCAGAGTGT 435
QY 201 ---AGCAGCTACTTAGCTGGTACCAACAGAACTGGCCAGGCTCCAGGCTCTCATC 257
Db 436 AGCAGCAGCTACTTAGCTGGTACCAACAGAACTGGCCAGGCTCCAGGCTCTCATC 495
QY 258 TATGATGCATCCACAGGGCCACTGGCATCCACCCAGGTTTCAGTGGCAGTGGTCTGG 317
Db 496 TATGATGCATCCACAGGGCCACTGGCATCCACCCAGGTTTCAGTGGCAGTGGTCTGG 555
QY 318 ACAGACTTCACTCTCACCATCAGCAGCTGGAGCCGGAAGATGTGGCAGCTTTTACTGT 377
Db 556 ACAGACTTCACTCTCACCATCAGCAGCTGGAGCCGGAAGATGTGGCAGCTTTTACTGT 615
QY 378 CAGCATATTTTA---CTACTCCGTACACTTTTGGCCAGGGAGCCAGGCTGGAGTCAAA 434
Db 616 CAGCATATTAATAACTGGGCTCGGTACACTTTTGGCCAGGGAGCCAGCTGGAGTCAAA 675
QY 435 CGAACTGTGGTGCACCATCTCTCTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCT 494
Db 676 CGAACTGTGGTGCACCATCTCTCTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCT 735
QY 495 GGAATGCTCTGTTGTGTGCTGCTGAATAACTTCTATCCAGAGGGCCAAAGTACAG 554
Db 736 GGAATGCTCTGTTGTGTGCTGCTGAATAACTTCTATCCAGAGGGCCAAAGTACAG 795
QY 555 TGAAGGTGGATAACGCCCTCCAACTCGGTAACCTCCAGGAGAGTGTACAGAGCAGGAC 614
Db 796 TGAAGGTGGATAACGCCCTCCAACTCGGTAACCTCCAGGAGAGTGTACAGAGCAGGAC 855
QY 615 ACAGAGGAGCAGCCTACAGCCTCAGCAGCACCCTGAGCTGAGCAAGCAGACTACGAG 674
Db 856 ACAGAGGAGCAGCCTACAGCCTCAGCAGCACCCTGAGCTGAGCAAGCAGACTACGAG 915
QY 675 AAACACAAAGTCTAGCCCTGCGAAGTCACCCATCAGGCTGAGCTGCGCCGTCACAAAG 734
Db 916 AAACACAAAGTCTAGCCCTGCGAAGTCACCCATCAGGCTGAGCTGCGCCGTCACAAAG 975
QY 735 AGCTTCAACAGGGAGAGTGTAGAGGAGAGTGGCCCGCCAGCTGCTCTCAGTTCCAGC 794
Db 976 AGCTTCAACAGGGAGAGTGTAGAGGAGAGTGGCCCGCCAGCTGCTCTCAGTTCCAGC 1035
QY 795 CTGACCCCTCCGATCTTTGGCTCTGACCCCTTTTCCAGAGGGAGCTACCCCTATTG 854
Db 1036 CTGACCCCTCCGATCTTTGGCTCTGACCCCTTTTCCAGAGGGAGCTACCCCTATTG 1095
QY 855 CGGTCTCCAGCTCATCTTTCACCTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 914
Db 1096 CGGTCTCCAGCTCATCTTTCACCTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 962
QY 915 ATGTTGAGGAGAGTAATAAATGTAATCTTTGCAAAAAA 962
Db 1156 ATGTTGAGGAGAGTAATAAATGTAATCTTTGCAAAAAA 1202

RESULT 5

US-09-822-830A-531

; Sequence 531, Application US/09822830A

Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fectel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 531
; LENGTH: 913
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-822-830A-531

Query Match 84.0%; Score 808.4; DB 11; Length 913;
Best Local Similarity 94.0%; Pred. No. 2.2e-226;
Matches 852; Conservative 0; Mismatches 51; Indels 3; Gaps 1;

QY 50 AACCATGGAAGCCCGAGCTCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 109
Db 1 AACCATGGAAGCCCGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
QY 110 CGGAGAAATGTTGTGACACAGTCTCCAGCCACCTCTCTCTCTCTCTCTCTCTCTCTCT 169
Db 61 CGGAGAAATGTTGTGACACAGTCTCCCGGCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 170 CACCCTCTCTCTCAGGCGCAGTCAGAG---TGTTAGCAGCTACTTAGCTTGGTACCAACA 226
Db 121 CACCCTCTCTCTAGGCGCAGTCAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 227 GAACTCTGCGCAGGCTCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 286
Db 181 GAACTCTGCGCAGGCTCCCGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 287 CCCACCCAGGTTCACTGGCAGTGGGTCTGGGACAGACTTCACCTCTCACCATCAGCAGACT 346
Db 241 CCCAGAAAGTTCACTGGCAGTGGGTCTGGGACAGACTTCACCTCTCACCATCAGCAGACT 300
QY 347 GGAGCCCGAAGATGTGGCAGTCTTATCTCTCAGCAATATTTTACTTACTCTCTCTCTCTCT 406
Db 301 GGAACCTGAAGACTTTTCGGGTATATTACTGTCTCACAATATCTAGTTCTCTCTCTCTCT 360
QY 407 TGGCCAGGGGACCAAGCTGGAGATCAACCAAGCTGGCTGGCAGCAGCTCTCTCTCTCTCTCT 466
Db 361 TGGCCAGGGGACCAAGCTGGTCTATCAAAAGCTGGCTGGCAGCAGCTCTCTCTCTCTCTCT 420
QY 467 CCGCCATCTGATGAGCAGTTGAAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAAT 526
Db 421 CCGCCATCTGATGAGCAGTTGAAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAAT 480
QY 527 CTCTATCTCCAGAGAGCCAAAGTACAGTGGAGGTGGATACAGCCCTCTCTCTCTCTCTCTCT 586
Db 481 CTCTATCTCCAGAGAGCCAAAGTACAGTGGAGGTGGATACAGCCCTCTCTCTCTCTCTCTCT 540
QY 587 CTCCAGGAGAGTGTCCAGAGCAGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 646
Db 541 CTCCAGGAGAGTGTCCAGAGCAGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 600
QY 647 CCTGACCTTGACCAAGCAGACTACGAGAAACAAAGTCTACGGCTCGGAAGTCAACCA 706
Db 601 CCTGACCTTGACCAAGCAGACTACGAGAAACAAAGTCTACGGCTCGGAAGTCAACCA 660

QY 707 TCAGGCGCTGAGCTGCGCGCTCACAAAGAGCTTCAACAGGGAGAGTGTAGAGGGAGAA 766
Db 661 TCAGGCGCTGAGCTGCGCGCTCACAAAGAGCTTCAACAGGGAGAGTGTAGAGGGAGAA 720
QY 767 GTGCGCCCACTGCTCTCAGTTCAGCTGACCCCTTCCAGCTCACTTTCACCTGACCC 826
Db 721 GTGCGCCCACTGCTCTCAGTTCAGCTGACCCCTTCCAGCTCACTTTCACCTGACCC 780
QY 827 TTTTTCACAGGGAGCTTACCCCTTATTCGCGTCTCCAGCTCACTTTCACCTGACCC 886
Db 781 TTTTTCACAGGGAGCTTACCCCTTATTCGCGTCTCCAGCTCACTTTCACCTGACCC 840
QY 887 CTCCTCTCTCTGCTTAAATATGCTGAGGAGAGTGAATAAAGTGAAT 946
Db 841 CTCCTCTCTCTGCTTAAATATGCTGAGGAGAGTGAATAAAGTGAAT 900
QY 947 CTTTGC 952
Db 901 CTTTGC 906

RESULT 6
US-09-954-456-771
; Sequence 771, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canceled
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIORITY APPLICATION NUMBER: US/60/233,617
; PRIORITY FILING DATE: 2000-09-18
; PRIORITY APPLICATION NUMBER: US/60/234,052
; PRIORITY FILING DATE: 2000-09-20
; PRIORITY APPLICATION NUMBER: US/60/234,923
; PRIORITY FILING DATE: 2000-09-25
; PRIORITY APPLICATION NUMBER: US/60/235,134
; PRIORITY FILING DATE: 2000-09-25
; PRIORITY APPLICATION NUMBER: US/60/235,637
; PRIORITY FILING DATE: 2000-09-26
; PRIORITY APPLICATION NUMBER: US/60/235,638
; PRIORITY FILING DATE: 2000-09-26
; PRIORITY APPLICATION NUMBER: US/60/235,711
; PRIORITY FILING DATE: 2000-09-27
; PRIORITY APPLICATION NUMBER: US/60/235,720
; PRIORITY FILING DATE: 2000-09-27
; PRIORITY APPLICATION NUMBER: US/60/235,840
; PRIORITY FILING DATE: 2000-09-27
; PRIORITY APPLICATION NUMBER: US/60/235,863
; PRIORITY FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 771
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-771

Query Match 82.4%; Score 792.4; DB 11; Length 1244;
Best Local Similarity 95.7%; Pred. No. 1.2e-221;
Matches 827; Conservative 0; Mismatches 31; Indels 6; Gaps 1;

QY 95 GCTCCAGATACCCAGGAGAAATGTGTACACAGTCTCCAGCCACCTGCTTTGTC 154
Db 381 GATCAACATACCCAGGAGAAATGTGTACACAGTCTCCAGCCACCTGCTTTGTC 440
QY 155 TCCAGGGGAAAGACCCCTCTCTCCAGGGCCAGTCAGAGTGTAGCAGCTACTTAC 214
Db 441 TCCAGGGGAAAGACCCCTCTCTCCAGGGCCAGTCAGAGTGTAGCAGCTACTTAC 500
QY 215 CTGTACCAACAGAAACCTGGCCAGGCTCCAGGCTCTCTATCTATGATGATCAACAG 274

Db 501 CTGCTACCAACAGAAACCTGGCCAGGCTCCAGGCGCCCTCATCTATGATCATCCACAG 560
QY 275 GGCCTACTGGCATCCCAACCCAGGTTTCAGTGGCAGTGGTCTGGGACAGACTTCACTCTCAC 334
Db 561 GGCCTACTGGCATCCCAACCCAGGTTTCAGTGGCAGTGGTCTGGGACAGACTTCACTCTCAC 620
QY 335 CATCAGCAGACTGGAGCCGGAAGATGTGGCAGTCTTATTTACTGTGACG-----AATATT 388
Db 621 CATCAGCAGCTAGAGCTGAAGATTTTGGAGTTTATTTACTGTCAACACCGCTGACAAATG 680
QY 389 TACTACTCCGTACACTTTTGGCCAGGGGACCAGGCTGGAGATCAAAAGCAACTGTGCTGCG 448
Db 681 GCTCCGCGGGCCCACTTTCGCGGAGGAGGACCAAGGTGGAGATCAAAAGCAACTGTGCTGCG 740
QY 449 ACCATCTGTCTTCATCTTCGCGGCACTGTATGAGCAGTTGAAATCTGGAACTGCTCTGT 508
Db 741 ACCATCTGTCTTCATCTTCGCGGCACTGTATGAGCAGTTGAAATCTGGAACTGCTCTGT 800
QY 509 TGTGTGCTGCTGAATAACTTTCTATCCAGAGAGGCGCAAGTACAGTGGAAAGTGGATAA 568
Db 801 TGTGTGCTGCTGAATAACTTTCTATCCAGAGAGGCGCAAGTACAGTGGAAAGTGGATAA 860
QY 569 GCGCTCCCAATCGGGTAACTCCAGGAGAGTGTACAGAGCAGGACAGCAAGGACAGCAC 628
Db 861 GCGCTCCCAATCGGGTAACTCCAGGAGAGTGTACAGAGCAGGACAGCAAGGACAGCAC 920
QY 629 CTACAGCCTCAGCAGCACCTGACGCTGAGCAAGCAGACTACGAGAAACAAAGTCTA 688
Db 921 CTACAGCCTCAGCAGCACCTGACGCTGAGCAAGCAGACTACGAGAAACAAAGTCTA 980
QY 689 CGCTGTGGAAGTCAACCCATCAGGGCTGAGCTGCGCCCTCACAAAGAGCTTCAACAGGG 748
Db 981 CGCTGTGGAAGTCAACCCATCAGGGCTGAGCTGCGCCCTCACAAAGAGCTTCAACAGGG 1040
QY 749 AGAGTGTAGAGGAGAGTGGCCCGCCCTGAGCTGCGCCCTCACAAAGAGCTTCAACAGGG 808
Db 1041 AGAGTGTAGAGGAGAGTGGCCCGCCCTGAGCTGCGCCCTCACAAAGAGCTTCAACAGGG 1100
QY 809 TCCTTTGGCTCTGACCCCTTTTCCACAGGGGAGCTACCCCTATTTCGCGTCTCCAGCTC 868
Db 1101 TCCTTTGGCTCTGACCCCTTTTCCACAGGGGAGCTACCCCTATTTCGCGTCTCCAGCTC 1160
QY 869 ATCTTTCAGCTCACCCCTCTCTCTCTGCTTTAATTAATGCTTAATGTTGGAGGAA 928
Db 1161 ATCTTTCAGCTCACCCCTCTCTCTCTGCTTTAATTAATGCTTAATGTTGGAGGAA 1220
QY 929 TGAATAAATAAAGTGAATCTTTTC 952
Db 1221 TGAATAAATAAAGTGAATCTTTTC 1244

RESULT 7

US-10-198-846-14013
; Sequence 14013, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIORITY APPLICATION NUMBER: 60/306,220
; PRIORITY FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14013
; LENGTH: 1001

[illegible]

Query Match 76.8%; Score 738.8; DB 10; Length 1033;
 Best Local Similarity 88.2%; Pred. No. 5.5e-206;
 Matches 822; Conservative 0; Mismatches 92; Indels 18; Gaps 1;

QY 49 GAACCTGAAGAGCCAGCTCAGCTTCTCTCCCTGCTACTCTGGCTCCAGATACCA 108
 Db 5 GCAAGATGTTGTCAGACCCAGGCTCTCAATCTCTGCTGCTGATCTCTGGTCCCT 64
 QY 109 CCGGAGAAATGTTGTGACAGTCTCCAGCACCTCTCTTTCTCTCCAGGGGAAGAG 168
 Db 65 ACGGCGACATCTGATGATCCAGCCAGTCTCCAGACTCCCTGGCTGTCTCTGGCGAGGG 124
 QY 169 CCACCTCTCTCCAGGCGCCAGTCAGAGTGT-----TAGCAGCTACT 210
 Db 125 CCACCATCACTGCAAGTCCAGCCAGAGCTGTTTATACAGCTCCGACAATAAGAACTACT 184
 QY 211 TAGCTGTGTACCAAGAAACCTGCGCAGGCTCCAGGCTCTCATCTATGATGCATCA 270
 Db 185 TAGCTGTGTACCAAGAAACCCAGACAGCTCTCAAGCTGCTCAATTTACTGGGCATCTA 244
 QY 271 ACAGGGCCACTGGCATCCACCCAGGTTCTAGTGCGAGTGTGGGACAGACTTTCACCTC 330
 Db 245 CCCGGGAATCCGGGTCCTGACCGATTCTAGTGCGACGGGTCTGGACAGATTTTCACCTC 304
 QY 331 TCACCATCAGCAGACTGGGACCCGAGGCTGGAGATCAACAACTGTGGCTGCAC 450
 Db 305 TCACCATCAGCAGCTGCAGCTGAAGATGTGCGAGTTTATTAATCTCAGCAATATTATA 364
 QY 391 CTACTCGGTACACTTTGGCCAGGGACCGAGGCTGGAGATCAACAACTGTGGCTGCAC 450
 Db 365 GTACTCGGTACAGTTTGGCCAGGGACCAAGCTGGAATCAACAACTGTGGCTGCAC 424
 QY 451 CATCTGTCTTCATCTCCGCCATCTGATGAGCAGTTGAATCTGGAAGTCTGCTGTGTG 510
 Db 425 CATCTGTCTTCATCTCCGCCATCTGATGAGCAGTTGAATCTGGAAGTCTGCTGTGTG 484
 QY 511 TGTGCTGTCTGAATACTTCTATCCAGAGAGGCCAAAGTACAGTGGAGGTGGATAAG 570
 Db 485 TGTGCTGTCTGAATACTTCTATCCAGAGAGGCCAAAGTACAGTGGAGGTGGATAAG 544
 QY 571 CCCTCCAAATCGGTTAACTCCAGGAGAGTGTACAGAGCAGGACAGCAAGGACACCT 630
 Db 545 CCCTCCAAATCGGTTAACTCCAGGAGAGTGTACAGAGCAGGACAGCAAGGACACCT 604
 QY 631 ACAGCTCAGCAGCACCCTGAGCTGAGCAAGCAGACTACGAGAAACACAAAGTCTACG 690
 Db 605 ACAGCTCAGCAGCACCCTGAGCTGAGCAAGCAGACTACGAGAAACACAAAGTCTACG 664
 QY 691 CTTGCGAAGTCAACCATCAGGCGCTGAGCTGCGCGGTCAACAAAGAGCTTCAACAGGGAG 750
 Db 665 CTTGCGAAGTCAACCATCAGGCGCTGAGCTGCGCGGTCAACAAAGAGCTTCAACAGGGAG 724
 QY 751 AGTGTAGAGGAGAGTGGCCGCCACCTGCTCCTCAGTTCCAGCTGACCCCTCCCATC 810
 Db 725 AGTGTAGAGGAGAGTGGCCGCCACCTGCTCCTCAGTTCCAGCTGACCCCTCCCATC 784
 QY 811 CTTTGGCTCTCAGCCCTTTTCCAGAGGGACCTACCCCTATTGCGGTCTCTCCAGCTCAT 870
 Db 785 CTTTGGCTCTCAGCCCTTTTCCAGAGGGACCTACCCCTATTGCGGTCTCTCCAGCTCAT 844
 QY 871 CTTTCACTCAGCCCT 920
 Db 845 CTTTCACTCAGCCCT 904
 QY 931 AATAAATAAGTGAATCTTTTGCACAAAAA 962
 Db 905 AATAAATAAGTGAATCTTTTGCACAAAAA 936

RESULT 11

US-10-158-646-59

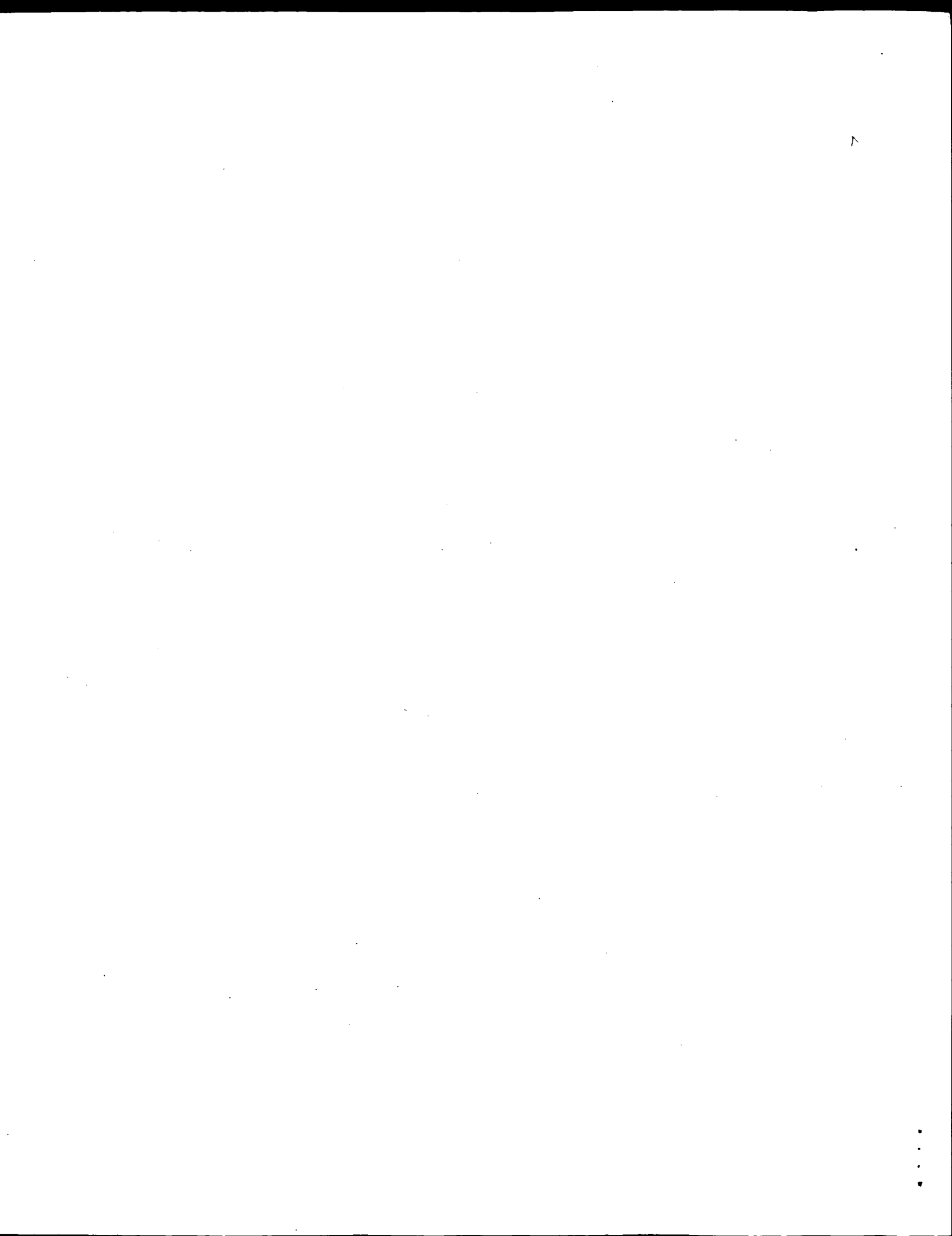
; Sequence 59, Application US/10158646

; Publication No. US20030073105A1

GENERAL INFORMATION:
 ; APPLICANT: Lasek, Amy K.W.
 ; APPLICANT: Sornasse, Thierry
 ; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
 ; FILE REFERENCE: PA-0030-1 US
 ; CURRENT APPLICATION NUMBER: US/10/158,646
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: 60/295,239
 ; PRIOR FILING DATE: 2001-05-31
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 59
 ; LENGTH: 1230
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: Incyte ID No. US20030073105A1 1101440.15
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: 1143-1160
 ; OTHER INFORMATION: a, t, c, g, or other
 US-10-158-646-59

Query Match 75.5%; Score 726.4; DB 15; Length 1230;
 Best Local Similarity 90.9%; Pred. No. 2.5e-202;
 Matches 802; Conservative 0; Mismatches 61; Indels 19; Gaps 2;

QY 100 CAGATACCCAGCGAGAAATTTGTTGACACAGTCTCCAGCCACCCTGT-CTTTGTCTCCA 158
 Db 249 CAGGTGCTACCGAGAAATTTGATGACCCAGTCTCCCTCCCTGGCTGTGTCTGTG 308
 QY 159 GGGGAAAGAGCCACCTCTCTGCGAGGCGCAGTGT-----T 200
 Db 309 GCGGAGAGGCGCACCACCTGAGGCGCCAGGAGTGTATACAGCTCCCAACAAT 368
 QY 201 AGCAGCTACTTACGCTGTGTACCAACAGAACTGCGCAGGCTCCAGGCTCTCATCTAT 260
 Db 369 AAGAATCTTGT 428
 QY 261 GATGATCAACAGGCGCCTGCGATCCACCCAGGTTTCTGCGAGTGTCTGGGACA 320
 Db 429 TGGGCTATCTACCGGGAATTCGGGTCCGACGACGATTCAGTGGCAGCGGTCTGGGACA 488
 QY 321 GACTTCACTCTCACCATCAGCAGTGTGAGCCCGGAGTGTGCGACCTTATTACTGTGAG 380
 Db 489 GATTTCACTCTCACCATCAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 548
 QY 381 CAATATTTTACTTCT 608
 Db 549 CAATATTTTACTTCT 668
 QY 441 GTGGCTGCACCATCTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 500
 Db 609 GTGGCTGCACCATCTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 668
 QY 501 GCCTCTGT 560
 Db 669 GCCTCTGT 728
 QY 561 GTGGATAACGCCCTTCAATCGGTAACTTCTATCCAGAGAGGCGCAAGTGTGAGTGTGAG 620
 Db 729 GTGGATAACGCCCTTCAATCGGTAACTTCTATCCAGAGAGTGTGAGTGTGAGTGTGAG 788
 QY 621 GACAGCCTTACAGCCTCAGCAGCCTGACCTGACCTGACCAAGAGAGTGTGAGTGTGAG 680
 Db 789 GACAGCCTTACAGCCTCAGCAGCCTGACCTGACCTGACCAAGAGAGTGTGAGTGTGAG 848
 QY 681 AAGTCTACGCTCGGAGTCAACCCATCAGGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 740
 Db 849 AAGTCTACGCTCGGAGTCAACCCATCAGGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 908
 QY 741 AACAGGGAGAGTGTGTAGAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 800



Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	829.6	86.2	931	3	US-09-049-672A-19	Sequence 19, Appl
2	751	78.1	3819	4	US-09-042-353-393	Sequence 393, App
3	751	78.1	3819	4	US-08-758-417A-243	Sequence 243, App
4	738.8	76.8	1066	1	US-08-157-101A-4	Sequence 4, Appl
5	685	71.2	958	3	US-09-049-672A-15	Sequence 15, Appl
6	580.2	60.3	8068	4	US-09-301-593-35	Sequence 35, Appl
7	565.2	58.8	13254	1	US-08-276-852-156	Sequence 156, App
8	565.2	58.8	13254	1	US-08-276-852-170	Sequence 170, App
c 9	565.2	58.8	13254	1	US-08-899-575-156	Sequence 156, App
c 10	565.2	58.8	13254	1	US-08-899-575-170	Sequence 170, App
c 11	565.2	58.8	13254	1	US-08-899-575-156	Sequence 156, App
c 12	565.2	58.8	13254	1	US-08-899-575-170	Sequence 170, App
c 13	565.2	58.8	13254	5	PCT-US95-08743-156	Sequence 156, App
c 14	565.2	58.8	13254	5	PCT-US95-08743-170	Sequence 170, App
c 15	563	58.5	729	1	US-08-276-852-152	Sequence 152, App
c 16	563	58.5	729	1	US-08-276-852-168	Sequence 168, App
c 17	563	58.5	729	1	US-08-899-575-152	Sequence 152, App
c 18	563	58.5	729	1	US-08-899-575-168	Sequence 168, App
c 19	563	58.5	729	1	US-08-899-575-152	Sequence 152, App
c 20	563	58.5	729	1	US-08-899-575-168	Sequence 168, App
c 21	563	58.5	729	5	PCT-US95-08743-152	Sequence 152, App
c 22	563	58.5	729	5	PCT-US95-08743-168	Sequence 168, App
c 23	553	57.5	4691	4	US-08-591-632-43	Sequence 43, Appl
c 24	553	57.5	4691	4	US-09-611-451-43	Sequence 43, Appl
c 25	553	57.5	6166	4	US-08-591-632-51	Sequence 51, Appl
c 26	553	57.5	6166	4	US-09-611-451-51	Sequence 51, Appl
c 27	551.4	57.3	646	1	US-08-300-386A-2	Sequence 2, Appl

CLONE: 2280869
US-09-049-672A-19

Query Match . 86.2%; Score 829.6; DB 3; Length 931;
Best Local Similarity 93.6%; Pred. NO. 3.9e-228;
Matches 865; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

```
QY 39 GACCCAGAGGGAACCATGAGAGCCCCAGCTCAGCTTCTTCCCTCCCTGCTACTCTGCTC 98
DB 1 GACCCAGAGGGAACCATGAGAGCCCCAGCTCAGCTTCTTCCCTCCCTGCTACTCTGCTC 60
QY 99 CCAGATACACCGGAGAAATGTGTGACACAGTCTCCAGCCACCCCTGTCTTTCTCTCCA 158
DB 61 CCAGATTCCTTTGGAGAAATGTAATGACGAGCTCTCCAGCCATCGTCTCTCTCTCCA 120
QY 159 GGGGAAGAGGACACCTCTCTGCGAGGCCAGTCAGAGTGTAGCAGTACTTACGCTGG 218
DB 121 GGGGAAGAGGACACCTCTCTGCGAGGCCAGTCAGTCTATTTGGCAGCACCATCGCTGG 180
QY 219 TACCAACAGAAACCTGGCCAGGCTCCAGGCTCTCTATGATGATGATGATGATGATGATG 278
DB 181 TACCAACAGAAACCTGGCCAGGCTCTCCAGGCTCTCTATGATGATGATGATGATGATG 240
QY 279 ACTGSCATCCCAACCCAGGTTGAGTGGCAGTGGGTCTGGGACAGACTTCACTCTCACCATC 338
DB 241 ACTGSGTCCCAACCCAGGTTGAGTGGCAGTGGGTCTGGGACAGAGTTCACTCTCTTCATC 300
QY 339 AGCAGACTGGAGCCCGAAGATGTGCACCTTTATTACTGTGACGAATATTTTACTACTCCG 398
DB 301 AGCAGCCTGCACTGGAAGATTTTGCACCTTTATTACTGTGACGAATATAAAGGTGGCGG 360
QY 399 TACACTTTTGGCCAGGGGACCCAGGCTGGAGATCAAAAGCTGGGTGACCACTGCTC 458
DB 361 CTCACCTTTGGGAGGAGGACCCAGGCTGGAGATCAAAAGCTGGGTGACCACTGCTC 420
QY 459 TTCATCTTCCGCCCATCTGATGAGCAGTTGAAATCTGGAATCTGCTCTGTGTGTGCTG 518
DB 421 TTCATCTTCCGCCCATCTGATGAGCAGTTGAAATCTGGAATCTGCTCTGTGTGTGCTG 480
QY 519 CTGAATACTCTTATCCAGAGAGGCAAAAGTACAGTGGAGGTGGAAGTAAAGCCCTCAA 578
DB 481 CTGAATACTCTTATCCAGAGAGGCAAAAGTACAGTGGAGGTGGAAGTAAAGCCCTCAA 540
QY 579 TCGGTAACTCCAGAGAGTGTACAGAGCAGGACAAAGGACAGCAGCCTACAGCCTC 638
DB 541 TCGGTAACTCCAGAGAGTGTACAGAGCAGGACAAAGGACAGCAGCCTACAGCCTC 600
QY 639 AGCAGACCTCAGCTGAGCAAGCAGACTACGAGAACACAAAGTCTAGCCTGCGAA 698
DB 601 AGCAGACCTCAGCTGAGCAAGCAGACTACGAGAACACAAAGTCTAGCCTGCGAA 660
QY 699 GTCACCCATCAGGCGCTGAGCTCGCCCGCTCAAAAGAGCTTCAACAGGGGAGAGTGTAG 758
DB 661 GTCACCCATCAGGCGCTGAGCTCGCCCGCTCAAAAGAGCTTCAACAGGGGAGAGTGTAG 720
QY 759 AGGAGAACTGCCCGACCTGCTCTCAGTCCAGCCTGACCCCTCCATCCTTTGGCC 818
DB 721 AGGAGAACTGCCCGACCTGCTCTCAGTCCAGCCTGACCCCTCCATCCTTTGGCC 780
QY 819 TCTGACCTTTTTCACAGGGGACCTACCCCTATTGCGGTCTCTCCAGCTCATCTTTCACC 878
DB 781 TCTGACCTTTTTCACAGGGGACCTACCCCTATTGCGGTCTCTCCAGCTCATCTTTCACC 840
QY 879 TCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 938
DB 841 TCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
QY 939 AAGTGAATCTTTCGCAAAAAAAA 962
DB 901 AAGTGAATCTTTCGCAAAAAAAA 924
```

RESULT 2

US-09-042-353-393
; Sequence 393, Application US/09042353
; Patent No. 6255458
; GENERAL INFORMATION:
; APPLICANT: Lonberg, Nils
; APPLICANT: Kay, Robert M.
; TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for
; TITLE OF INVENTION: Producing Heterologous Antibodies
; NUMBER OF SEQUENCES: 421
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,353
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/810,279
; FILING DATE: 17-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/853,408
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/904,068
; FILING DATE: 23-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,131
; FILING DATE: 26-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/096,762
; FILING DATE: 22-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/155,301
; FILING DATE: 18-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,739
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/165,699
; FILING DATE: 10-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/209,741
; FILING DATE: 09-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/352,322
; FILING DATE: 07-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/544,404
; FILING DATE: 10-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/728,463
; FILING DATE: 10-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US96/16433
; FILING DATE: 10-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,417
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/21803
; FILING DATE: 01-DEC-1997

us-09-806-276a-3.rni

Tue Jul 29 09:34:11 2003

```

413 GGGGACAGGCTGGAGATCAAAACAGAACTGTGGCTGACCATCTCTCTCTCTCTCTCGGCC 472
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
398 AGGACACAGGCTGGAGATCAAAACAGAACTGTGGCTGACCATCTCTCTCTCTCTCTCGGCC 457
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
473 ATCTGATGACAGTGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGA 532
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
458 ATCTGATGACAGTGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGA 517
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
533 TCCGAGAGAGGCGCAAAAGTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 592
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
518 TCCGAGAGAGGCGCAAAAGTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 577
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
593 GGAGAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 652
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
578 GGAGAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 637
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
653 GCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 712
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
638 GCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 697
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
713 CTTGAGCTGCGCCGCTCAAAAGTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 772
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
698 CTTGAGCTGCGCCGCTCAAAAGTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 757
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
773 CCACCTGCTCTCAGTCTCCAGCTGAGCCCTCCAGCTGAGCCCTCCAGCTGAGCCCTCCAGCTGAGCC 832
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
758 CCACCTGCTCTCAGTCTCCAGCTGAGCCCTCCAGCTGAGCCCTCCAGCTGAGCCCTCCAGCTGAGCC 817
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
833 CACAGGGGACCTTACCCCTATTGCGGTCTCCAGCTGAGCCCTCCAGCTGAGCCCTCCAGCTGAGCC 892
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
818 CACAGGGGACCTTACCCCTATTGCGGTCTCCAGCTGAGCCCTCCAGCTGAGCCCTCCAGCTGAGCC 877
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
893 CTCTTGGCTTAAATATGCTAATGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 952
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
878 CTCTTGGCTTAAATATGCTAATGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 937
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
953 AAAAAAAAAA 962
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
938 AAAAAAAAAA 947
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

RESULT 5
US-09-049-672A-15
; Sequence 15, Application US/09049672A
; Patent No. 6135941
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Kal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yue, Henry
; APPLICANT: Au-Young, Janice
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/049,672A
; FILING DATE: HEREWITH
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ceirone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0497 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: ADENINB01
CLONE: 161752
US-09-049-672A-15

Query Match 71.2%; Score 685; DB 3; Length 958;
Best Local Similarity 84.6%; Pred. No. 1e-186;
Matches 769; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

413 GGGGACAGGCTGGAGATCAAAACAGAACTGTGGCTGACCATCTCTCTCTCTCTCTCGGCC 113
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
43 ATGGGGTCCCAGGTTCCACCTCCCTCAGCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 102
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
114 GAAATGTGTGACACAGCTCCAGCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 173
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
103 GAAACGACACTACAGGAGTCTCCAGTATTCATGTCTCAGGACTCCAGGAGAGACAAAGTCAAC 162
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
174 CTCTCTCTGAGGGCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 233
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
163 ATCTCTGCAAGCCAGCCAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 222
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
234 GGCAGGCTCCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 293
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
223 GGAGAAGTGTATTTTTCATATTTTTCATATTTTTCATATTTTTCATATTTTTCATATTTTTCATAT 282
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
294 AGGTTTCAAGTGGAGTGTGGGACAGAGTTCACCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 353
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
283 CGATTCAAGTGGAGTGTGGGACAGAGTTCACCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 342
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
354 GAAGATGAGGACCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 413
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
343 GAGGATCTGATATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 402
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
414 GGGACAGGCTGGAGATCAAAACAGAACTGTGGCTGACCATCTCTCTCTCTCTCTCTCTCTCTCTCT 473
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
403 GGGACAGGCTGGAGATCAAAACAGAACTGTGGCTGACCATCTCTCTCTCTCTCTCTCTCTCTCTCT 462
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
474 TCTGATGAGCAGTTGAAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGA 533
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
463 TCTGATGAGCAGTTGAAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGA 522
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
534 CCCAGAGAGGCGCAAAAGTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 593
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
523 CCCAGAGAGGCGCAAAAGTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 582
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
594 GAGAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 653
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
583 GAGAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 642
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
654 CTGAGCAAAAGCAGACTACGAGAAACACAAAGTCTACGCTCTGGAAGTCAACCCATCAGGCT 713
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
643 CTGAGCAAAAGCAGACTACGAGAAACACAAAGTCTACGCTCTGGAAGTCAACCCATCAGGCT 702
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
714 CTGAGCTCGCCGCTCACAAGAGCTTCAACAGGGGAGAGTGTAGAGGGAGAGAGTGGCC 773
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
703 CTGAGCTCGCCGCTCACAAGAGCTTCAACAGGGGAGAGTGTAGAGGGAGAGTGGCC 762
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

QY	774	CACCTGCTCCTCAGTTCCAGCCTCACC	CCCTCCATCCTTTGGCCCTCTGAC	CCCTTTTCC	833
Db	763	CACCTGCTCCTCAGTTCCAGCCTCACC	CCCTCCATCCTTTGGCCCTCTGAC	CCCTTTTCC	822
QY	834	ACAGGGGACCTACCCCTATTGCGGTCT	CCTCCAGCTCATCTTTCACCTCAC	CCCGCTCTCTCC	893
Db	823	ACAGGGGACCTACCCCTATTGCGGTCT	CCTCCAGCTCATCTTTCACCTCAC	CCCGCTCTCTCC	882
QY	894	TCCTTGGCTTTAAATTATGCTAAATG	TTGGAGGAGATGAATAAATAAGAT	CTTTGCA	953
Db	883	TCCTTGGCTTTAAATTATGCTAAATG	TTGGAGGAGATGAATAAATAAGAT	CTTTGCA	942
QY	954	AAAAAAAAA	962		
Db	943	AAAAAAAAA	951		

RESULT 6

US-09-301-593-35
: Sequence 35, Application US/09301593A
: Patent No. 6455677
: GENERAL INFORMATION:
: APPLICANT: Park, John E.
: APPLICANT: Garin-Chesa, Pilar
: APPLICANT: Bamberger, Uwe
: APPLICANT: Legeer, Olivier
: APPLICANT: Saidanha, Jose W.
: APPLICANT: Rettig, Wolfgang J.
TITLE OF INVENTION: FAP-specific Antibody with Improved Producibility
FILE REFERENCE: 0652.1890001
CURRENT APPLICATION NUMBER: US/09/301,593A
CURRENT FILING DATE: 1999-04-29
EARLIER APPLICATION NUMBER: EP 98107925.4
EARLIER FILING DATE: 1998-04-30
EARLIER APPLICATION NUMBER: US 60/086,049
EARLIER FILING DATE: 1998-05-18
NUMBER OF SEQ ID NOS: 108
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 35
LENGTH: 8068
TYPE: DNA
ORGANISM: Homo sapiens
S-09-301-593-35

Query Match	60.3%;	Score 580.2;	DB 4;	Length 8068;
Best Local Similarity	76.2%;	Pred. No. 2.8e-156;		
Matches 813;	Conservative 0;	Mismatches 103;	Indels 151;	Gaps 2;

QY	38	GGACCCAGAGGGAACCAATGGTGAAGCCGCCAGCTCAGCTTCCTCTTCCTCCCTGCTACTCTGGCT	97
Db	1583	GAAGCTTGGCCGCCACCATGAGACACACACTCCCTGCTATGGGTGCTGCTCTGGT	1642
QY	98	CCAGATACCAACGGGAATATGTGTGACAGACTCTCCAGCCACCGCTGTCTTGTCCTCC	157
Db	1643	TCCAGGTTCCTCGGAGACATTTGTGATGACCCAATCTCCAGACTCTTTGGCTGTGCTCT	1702
QY	158	AGGGAAAGAGCCACCCTCTCTGCAGGCGCAGTCAGAGTCTT	200
Db	1703	AGGGAGAGGCGCCACCATCACTGCAAGTCCAGTCAGAGCCTTTTATATCTAGAAATCA	1762
QY	201	AGCAGCTACTTAGCCCTGGTACCAACAGAAACCTGGCCAGCTCCCAAGGCTCTCATCTA	259
Db	1763	AAAGAACTACTTTGGCCTGTATCACAGAAACCAAGGACACCCAACTCCTCATCTT	1822
QY	260	TGATGATCAACAGGCGCACTGGCATCCCAACCCAGGTTTCAGTGGCAGTGGGTCTGGGAC	319
Db	1823	TTGGGCTAGCATAGGGAATCTGGGTAACCTTGATAGGTTTCAGTGGCAGTGGGTGGGAC	1882
QY	320	AGACTTCACCTCACCATCAGCAGACTGGAGCCGAGATGTGGCACTTTATTACTGTCA	379
Db	1883	AGACTTCACCTCACCATTAGCAGCCCTGCAGGCTGAAGATGTGGCACTTTATTACTGTCA	1942
QY	380	GCAATATTTTACTACTCGGTACACTTTTGGCCAGGGGACACCGCTGGAGATCAAAAC	435

Db	1943	GCATATTTTAGCTATCCGCTCAGGTTTCGGACAAGGCACCAAGGTGGAAATAAAACGTGA	2002
QY	436	-----	435
Db	2003	GTGGATCCATCTGGGATAAGCATGCTGTTTCTGTGTCGTCCCTAACATGCCCCGTGTGATTA	2062
QY	436	-----	435
Db	2063	TGCGCAAAACACACACCCAAAGGCGCAAACTTTGTTACTTAAACACCATCCTGTTTGCCTC	2122
QY	436	-----GAACCTGTGGCTGCAACCATCTGCTCTTTCATCTTCCGGCCATCTGATGAGCAGT	486
Db	2123	TTTCCTCAGGAACTGTGGCTGCAACCATCTGCTCTCATCTTCCGGCCATCTGATGAGCAGT	2182
QY	487	TGAAATCTGAACTCGCTCTGTTGTGTGCTGCTGCTGCTGCTGAATAACTTCTATCCCAGAGAGGCCA	546
Db	2183	TGAAATCTGAACTCGCTCTGTTGTGTGCTGCTGCTGAATAACTTCTATCCCAGAGAGGCCA	2242
QY	547	AAGTACAGTGGAAAGTGGGATAACGCCCTCCAACTCGGGTAACTPCCAGAGAGAGTGTCCACAG	606
Db	2243	AAGTACAGTGGAAAGTGGGATAACGCCCTCCAACTCGGGTAACTPCCAGAGAGAGTGTCCACAG	2302
QY	607	AGCAGACACAGCAAGGACACACCTACAGCCTCAGCAGCACCCCTGACGCTGAGCAAAAGCAG	666
Db	2303	AGCAGGACACAGCAAGGACACACCTACAGCCTCAGCAGCACCCCTGACGCTGAGCAAAAGCAG	2362
QY	667	ACTACAGAGAAACACAAAGTCTACGCCCTCGGAAGTCACCCATCAGGGCTGAGCTCGCCCG	726
Db	2363	ACTACAGAGAAACACAAAGTCTACGCCCTCGGAAGTCACCCATCAGGGCTGAGCTCGCCCG	2422
QY	727	TCACAAAGAGCTTCAACAGGGGAGAGTGTAGAGGGGAGAAGTGCCTCCCTCCTCTCA	786
Db	2423	TCACAAAGAGCTTCAACAGGGGAGAGTGTAGAGGGGAGAAGTGCCTCCCTCCTCTCA	2482
QY	787	GTTCAGGCTGACCCCTCCCATCCTTTGGGCTCTGACCCCTTTTCCACAGGGGACCTAC	846
Db	2483	GTTCAGGCTGACCCCTCCCATCCTTTGGGCTCTGACCCCTTTTCCACAGGGGACCTAC	2542
QY	847	CCCTATTGCGGTCTCCAGGCTCATCTTTACCTCACCCCCCTCCTCTCTTGGGCTTTAA	906
Db	2543	CCCTATTGCGGTCTCCAGGCTCATCTTTACCTCACCCCCCTCCTCTCTTGGGCTTTAA	2602
QY	907	TTATGCTTAATGTTGGAGGAGAAATGAATAAATAAGTGAATCTTTGCA	953
Db	2603	TTATGCTTAATGTTGGAGGAGAAATGAATAAATAAGTGAATCTTTGCA	2649

RESULT. T 7

US-08-276-852-156
Sequence 156, Application US/08276852
Patent No. 5652138
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25

/	FILING DATE:	24-JUL-1997
/	CLASSIFICATION:	
/	PRIOR APPLICATION DATA:	
/	APPLICATION NUMBER:	US 08/276,852
/	FILING DATE:	18-JUL-1994
/	APPLICATION NUMBER:	US 08/178,302
/	FILING DATE:	30-SEP-1993
/	PRIOR APPLICATION DATA:	
/	APPLICATION NUMBER:	US 07/954,148
/	FILING DATE:	30-SEP-1992
/	ATTORNEY/AGENT INFORMATION:	
/	NAME:	Fitting, Thomas
/	REGISTRATION NUMBER:	34,163
/	REFERENCE/DOCKET NUMBER:	SCRI452P
/	TELECOMMUNICATION INFORMATION:	
/	TELEPHONE:	619-554-2937
/	TELEFAX:	619-554-6312
/	INFORMATION FOR SEQ ID NO:	156:
/	SEQUENCE CHARACTERISTICS:	
/	LENGTH:	13254 base pairs
/	TYPE:	nucleic acid
/	STRANDEDNESS:	double
/	TOPOLOGY:	circular
/	MOLECULE TYPE:	DNA (genomic)
/	US-08-899-575-156	
Query Match		
Best Local Similarity		
Matches 634; Conservative 0; Mismatches 93; Indels 3; Gaps 1;		
QY	39	GACCCAGAGGAACCATGGAAAGCCCACGCTCAGCTTCTCTTCCTCGTCTACTCTGGCTC 98
DB	12522	GACACGAAGCTTACCATGGGTGTGCCACTCAGGTCCTGGGGTTTGCTGCTGTGTGGCTT 12581
QY	99	CCGATACCAACCCGAGAAATTGTTGCACAGTCTCCAGCACCOCTGCTTTGTCTCCA 158
DB	12582	ACAGATGCCAGATGTGAGATCGTCTCACGCAGTCTCCAGGCACCOCTGCTCTGTCTCCA 12641
QY	159	GGGGAAGAGCCACCCTCTCCTCAGGGCCAGTTCAGAGTGTT---AGCAGTACTATTAGCC 215
DB	12642	GGGGAAGAGCCACCCTCTCCTGTAGTCCAGTCACAGATTCGACGCCGGCGGTAGCC 12701
QY	216	TGTTACCAACAGAAAACCTGGCCAGGCTCCAGGCTCTCATCTAFTAGTCATCCAACAGG 275
DB	12702	TGTTACCAACAGAAAACCTGGCCAGGCTCCAGGCTGGTCATACATGTTGTTTCCAATAGG 12761
QY	276	GCCACTGGCATCCCCACCCAGGTTCACTGGCAGTGGGTCTGGGACAGACTTCACCTCACC 335
DB	12762	GCCTCTGGCATCTCAGACAGGTTACGGSCAGTGGGTCTGGACAGACTTCACCTCTCACC 12821
QY	336	ATCAGCAGACTGGAGCCCGAAGATGCGCACTTTATTACTGTACAATAATTTTTACTACT 395
DB	12822	ATCACCAGAGTGGAGCCTGAAGACTTTGCACCTGTACTACTGTCAAGTCTATGTTGCCTCC 12881
QY	396	CCGTACACTTTTGGCCAGGGACCAAGGCTGGAGATCAAACCAACTGTGGCTGCACCACT 455
DB	12882	TGCTACACTTTTGGCCAGGGACCAAACTGGAGAGAAACCAACTGTGGCTGCACCACT 12941
QY	456	GTCTTCATCTCCCGCCATCTGATGAGCAGTTGAAATCTGGAATCTGGAATCTGTTGTGTC 515
DB	12942	GTCTTCATCTCCCGCCATCTGATGAGCAGTTGAAATCTGGAATCTGGAATCTGTTGTGTC 13001
QY	516	CTGCTGAATAACTTCTATCCAGAGAGCCAAAGTACAGTGAAGAGTGGATTAACGCCCTC 575
DB	13002	CTGCTGAATAACTTCTATCCAGAGAGCCAAAGTACAGTGAAGAGTGGATTAACGCCCTC 13061
QY	576	CAATCGGGTAACCTCCAGGAGAGTGTACAGAGCAGGACAGCAAGAGCAGACCTACAGC 635
DB	13062	CAATCGGGTAACCTCCAGGAGAGTGTACAGAGCAGGACAGCAAGAGCAGACCTACAGC 13121
QY	636	CTCAGCAGCCOCTCAGCCTGAGCAAGCAGACTTACAGAGAAACACAAAGTCTACGCCGTC 695
DB	13122	CTCAGCAGCCOCTCAGCCTGAGCAAGCAGACTTACAGAGAAACACAAAGTCTACGCCGTC 13181

PCT-US95-08743-156					
	Query Match	58.8%;	Score 565.2;	DB 5;	Length 13254;
	Best Local Similarity	86.8%;	Pred. No. 6.8e-152;		
	Matches 634;	Conservative	0; Mismatches 93;	Indels	Gaps 1;
QY	39	GACCCAGAGGGAAACCATTGGAAGGCCCAAGCTCAGCTTCTCTTCCTCTCGTACTCTGTGGCTC	98		
Db	12522	GACAGAAGCTTTACCATGGGTGTGCCACTCAGGTCTGGGTTGCTGCTGTGTGGCTT	12581		
QY	99	CCAGATACCACCGGAGAAAATTGTCTTGACACAGCTCTCCAGCACACCTGTCTTTGTCTCCA	158		
Db	12582	ACAGATGCCAGATGTGAGATCGTTCTACGCCAGTCTCCAGGCACCTGTCTGTCTCCA	13641		
QY	159	GGGGAAGAGCACCCTTCTCTCGAGGGCAGTCAGAGTGT ---AGCAGCTACTTAGCC	215		
Db	12642	GGGNAAGAGCCACTTCTCTGTAGTGCAGTCACAGATTTCGCAGCGCGCGGTAGCC	12701		
QY	216	TGTTACCAACAAGAACCTTGCCAGGCTCCAGGCTCTCATCTATGATGCATNCAACAGG	275		
Db	12702	TGTTACCAAGACAAACCTTGSCCAGGCTCCAGGCTGTGCATACATGGTGTTCATAAG	12761		
QY	276	GCCACTGGCATCCCACCCAGGTTCACTGCGAGTGGGTCTGGGACAGACTTCACCTCTCAAC	335		
Db	12762	GCCTCTGGCATCTCAGCAGGTTTCACGGGCTACAGCGGCTGGGTCTGGGACAGACTTCACCTCAAC	12821		
QY	336	ATCAGCAGCTGGAGCCCCAAGATGTGGCACTTTATTACTCTCAGCAATATTTACTACT	395		
Db	12822	ATCACCAGAGTGGAGCCTGAAGACTTTGCACCTGTACTACTGTGAGGTCTATGGTGGCTCC	12881		
QY	396	CCGTACACTTTTGGCCAGGGACAGGCTGGAGATCAAACGAACCTGTGGCTGCACACTC	455		
Db	12882	TCGTACACTTTTGGCCAGGGACCAACTGGAGAGAAACGAACCTGTGCTGCACACTC	12941		
QY	456	GTCTTCATCTCCGCCCATCTGATGAGCAGTTGAATCTGGAACTGCCCTCTGTTGTGTC	515		
Db	12942	GTCTTCATCTTCCGCCCATCTGATGAGCAGTTGAATCTGGGACTGCCCTCTGTTGTGTC	13001		
QY	516	CTGCTGAATAACTCTATCCACAGAGGCCAAAGTACAGTGGAGGTGGATAACGCCCTC	575		
Db	13002	CTGCTGAATAACTCTATCCACAGAGGCCAAAGTACAGTGGAGGTGGATAACGCCCTC	13061		
QY	576	CAATCGGTAACTCCACAGGAGTGTACAGACAGGACGACGACGACGACGACCTACAGC	635		
Db	13062	CAATCGGTAACTCCACAGGAGTGTACAGACAGGACGACGACGACGACGACCTACAGC	13121		
QY	636	CTCAGCAGCACCTTGACGCTTGACCAAGACAGACTACGAGAACACAAAGTCTACGCTGC	695		
Db	13122	CTCAGCAGCACCTTGACGCTTGACCAAGACAGACTACGAGAACACAAAGTCTACGCTGC	13181		
QY	696	GAAGTCACCATCAGGCTTGAGCTCGCCGTCACAAAGACTTCAACAGGGGAGAGTGT	755		
Db	13182	GAAGTCACCATCAGGCTTGAGCTCGCCGTCACAAAGACTTCAACAGGGGAGAGTGT	13241		
QY	756	TAGAGGAGA	765		
Db	13242	TAATCTTGA	13251		

RESULT 14
 PCT-US95-08743-170/G
 : Sequence 170, Application PC/TUS9508743
 : GENERAL INFORMATION:
 : APPLICANT:
 : TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
 : TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
 : NUMBER OF SEQUENCES: 170
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 : CURRENT APPLICATION DATA:

Sequence 152, Application US/08276852
Patent No. 5652138
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TO HUMAN IMMUNODEFICIENCY VIRUS
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESS: Patent Counsel
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA: US 07/954,148
APPLICATION NUMBER: 30-SEP-1992
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 729 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 9..715
US-08-276-852-152

Query Match	58.5%;	Score 563;	DB 1;	Length 729;
Best Local Similarity	87.1%;	Pred. No. 7.9e-152;	Indels	3; Gaps
Matches 630;	Conservative	0;	Mismatches 90;	
46	AGGACCATGGAAGCCCGAGCTCAGCTCTCTTCCTCGTGCTACTGTGCTCCCGACATA	105		
1	AGCTTCCATGGGTGGCCACTCAGGTCCTGGGGTTGCTGCTGTGGCTGTACAGATG	60		
106	CCACCGGAGAAATGTTGTACACAGCTCCAGGCACCCCTGCTTTGTCTCCAGGGGAAA	165		
61	CCAGATGTGAGATCGGTTCTACGCAGTCTCCAGGCACCCCTGCTCTGTCTCCAGGGGAAA	120		
166	GAGCCACCCCTCTCCTCGAGGGCCAGTCAGAGTGTT---AGCAGCTACTTACGCTGGTACC	222		
121	GAGCCACTTCTCTGTGTAGTCCAGTCACAGATTCGACGCCGCGGTAGCCTGGTACC	180		
223	AACAGAACCTTGGCCAGGCTCCAGGCTCCTCATATATGATGCATCCACAGGGCCACTG	282		
181	AGCACAACTTGGCCAGGCTCCAAAGGCTGGTCTCATCATATGTTTCCCAATAGGGCCTCTG	240		
283	GCATCCCAACCGATTTCAGTGGCAGTGGGTCTGGGACAGACTTCATCTCACCATCAGCA	342		

APPLICATION NUMBER: PCT/US95/08743
FILING DATE: 11-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRAINEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
PCT/US95-08743-170

	Query Match	58.8%;	Score 565.2;	DB 5;	Length 13234;
	Best Local Similarity	86.8%;	Pred. No. 6.8e-152;		
	Matches 634;	Conservative	0; Mismatches 93;	Indels	Gaps 1;
QY	39	GAACAGAGGGAAACCATGGAAGCCCGCAGCTTCCTTCTCTCTCTACTCTGGCTC	98		
Db					
QY	733	GACACGAAGCTTACCATTGGGTGTGCCCACTCAGGTCTCTGGGGTGCTGCTGTGGCTT	674		
Db					
QY	99	CCAGATACCACCGGAGAAATGTGTGTACACAGTCTCCAGCCACCTGTCTTTGTCTCCA	158		
Db					
QY	673	ACAATGCCAGATGTGAGATCGTTCTCACGCAGTCTCCAGGCACCTGTCTGTCTCCA	614		
Db					
QY	159	GGGAAAAGAGCCACCTCTCTCTGAGGGCCAGTCAGAGTCTTT---AGCAGCTACTTAGCC	215		
Db					
QY	613	GGGAAAAGAGCCACCTCTCTCTGTAGTCCAGTTCACAGCATTCGACGCCGCCGTAGCC	554		
Db					
QY	216	TGGTACCACAGA AACCTGCCAGGCTCCAGGCTCCCTCATCTATGATGATCCACAGG	275		
Db					
QY	553	TGGTACCAGCACAACTTGCCAGGCTCCAGGCTGGTTCATACATGGTGTTCCAATAGG	494		
Db					
QY	276	GCACCTGGCATCCCA CCCCAGGTTCAGTGCAGCTGGGTCTGGGACAGACTTCACTCTCACC	335		
Db					
QY	493	GCCTCTGGCATCTCAGACAGAGGTTTCAGGGCGATGGGTCTGGGACAGACTTCACTCTCACC	434		
Db					
QY	336	ATCAGCAGACTGGAGCCGCGAAGATGTGGCACTTTATTACTGTACAGCAATTTTACTACT	395		
Db					
QY	433	ATCCAGAGTGGAGCCTGAAAGACTTTGCACCTGTACTACTGTCAAGTCTATGGTGGCTCC	374		
Db					
QY	396	CCGTACACTTTTGGCCAGGGACCA GCGCTGGAGATCAAACGAACCTGTGGCTGCACCACT	455		
Db					
QY	373	TCGTACACTTTTTGGCCAGGGGACCA AACTCGAGAGGAACGAACCTGCGCTGCACCACT	314		
Db					
QY	456	GTCTTCACTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACCTCCCTCTGTGTGTGC	515		
Db					
QY	313	GTCTTCNTCTCCCGCCATCTGATGAGCAGTTGAAATCTGGGACTGCGCTGTGTGTGTGC	254		
Db					
QY	516	CTGCTGAATTA ACTTTATCCAGAGAGGCCAAAGTACAGTGGAAAGTGGAATAAGCCCTC	575		
Db					
QY	253	CTGCTGAATTA ACTTTATCCAGAGAGGCCAAAGTACAGTGGAAAGTGGAATAAGCCCTC	194		
Db					
QY	576	CAATCGGCTAACTCCCA GAGAGAGTGTACAGAGCAGGACAGAACGACAGCACTTACAG	635		
Db					
QY	193	CAATCGGCTAACTCCCA GAGAGAGTGTACAGAGCAGGACAGAACGACAGCACTTACAG	134		
Db					
QY	636	CTCAGCAGCACCTTGAGCTTGAGCAACGACACTACGAGAAACACAAAAGTCTACGCCCTGC	695		
Db					
QY	133	CTCAGCAGCACCTTGAGCTTGAGCAACGACACTACGAGAAACACAAAAGTCTACGCCCTGC	74		
Db					
QY	696	GAACTACCCCATCAGGSCCTTGAGCTCGCCCGCTCA AAAAGCTTCAACAGGGGAGAGTGT	755		
Db					
QY	73	GAACTACCCCATCAGGSCCTTGAGCTCGCCCGCTCA AAAAGCTTCAACAGGGGAGAGTGT	14		
Db					
QY	756	TAGAGGGAGA	765		
Db					
QY	13	TAATTCAGA	4		

RESULT 15
US-08-276-852-152

Search completed: July 25, 2003, 19:56:46
Job time : 51.6192 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 14:13:55 ; Search time 2529.23 Seconds
(without alignments)
11069.353 Million cell updates/sec

Title: US-09-806-276A-3
Perfect score: 962
Sequence: 1 tcgaccgattcgctcgag.....gaattttgcaaaaaaaaaa 962

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	870	90.4	1020	9	BC030813	BC030813 Homo sapi
2	857	89.1	968	9	BC016380	BC016380 Homo sapi
3	845.4	87.9	970	6	AX305008	AX305008 Sequence
4	845.4	87.9	970	6	AX305008	AX305008 Sequence
5	836.8	87.0	997	9	BC032451	BC032451 Homo sapi
6	830.2	86.3	948	6	AX305004	AX305004 Sequence
7	830.2	86.3	948	6	AX305004	AX305004 Sequence
8	823.6	86.2	931	6	AR135361	AR135361 Sequence
9	823.2	85.6	924	6	E07333	E07333 cDNA sequen
10	792.4	82.4	1244	6	AX333289	AX333289 Sequence
11	792.4	82.4	1244	9	HUMIGK	M33438 Human 1g re
12	756.4	78.6	956	9	BC029444	BC029444 Homo sapi
13	753.2	78.3	962	9	BC034141	BC034141 Homo sapi
14	751	78.1	3819	6	AR161402	AR161402 Sequence
15	746.8	77.6	953	9	BC005332	BC005332 Homo sapi
16	746.8	77.6	974	6	AX305000	AX305000 Sequence
17	746.8	77.6	974	6	AX305000	AX305000 Sequence
18	740.4	77.0	938	6	AX067344	AX067344 Homo sapi
19	735.6	76.5	944	6	AX067344	AX067344 Sequence
20	734	76.3	938	9	HSIGGKL	Y14736 Homo sapien
21	734	76.3	961	9	BC034146	BC034146 Homo sapi
22	726	75.5	944	9	BC018761	BC018761 Homo sapi
23	716.4	74.5	979	9	BC017870	BC017870 Homo sapi
24	710.4	73.8	928	6	AX251598	AX251598 Sequence
25	704.8	73.3	929	9	AK097353	AK097353 Homo sapi
26	701.2	72.9	977	9	BC034142	BC034142 Homo sapi
27	695	72.2	972	9	BC030814	BC030814 Homo sapi
28	691.8	71.9	973	9	BC022362	BC022362 Sequence
29	685	71.2	958	6	AR135357	AR135357 Sequence
30	683.6	71.1	802	6	AX067324	AX067324 Sequence
31	683.2	71.0	1952	9	AK024974	AK024974 Homo sapi
32	680	70.7	945	6	A21386	A21386 Plasmid DNA
33	680	70.7	963	6	AX067347	AX067347 Sequence
34	678.6	70.5	944	9	AK096938	AK096938 Homo sapi
35	656.8	68.3	708	9	AX419497	AX419497 Sequence
36	615.6	64.0	708	9	AF027158	AF027158 Homo sapi
37	601	62.5	811	6	AB064068	AB064068 Homo sapi
38	600	62.4	891	6	E40748	E40748 Antihuman F
39	599.4	62.3	830	9	AB064109	AB064109 Homo sapi
40	599.2	62.3	794	9	AB064129	AB064129 Homo sapi
41	599	62.3	908	9	AF186176	AF186176 Homo sapi
42	596	62.0	781	9	AB064131	AB064131 Homo sapi
43	592.8	61.6	807	9	AB064071	AB064071 Homo sapi
44	592.6	61.6	833	9	AB064110	AB064110 Homo sapi
45	591.4	61.5	738	9	AB064124	AB064124 Homo sapi

ALIGNMENTS

RESULT 1
BC030813
LOCUS BC030813 1020 bp mRNA linear PRI 13-JUN-2002
DEFINITION Homo sapiens, clone MGC:22645 IMAGE:4700961, mRNA, complete cds.
ACCESSION BC030813
VERSION BC030813.1 GI:21410095
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1020)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2002) National Institutes of Health, Mammalian

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 38 Row: j Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

```

FEATURES
source
Location/Qualifiers
1..968
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:27376 IMAGE:468477"
/tissue_type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
38..745
/codon_start=1
CDS

```

```

/product="similar to immunoglobulin kappa constant."
/protein_id="AAH16380.1"
/db_xref="GI:16741061"
/translation="METPAQLFLLLLLPGLTGEIVLTQSPATLSLSPGERATLSC
QAVSYSAWYQKQQAQLFLMGSSRAATGIPFRRSGSGDTLTLLSLQRE
DFAVYCYQGSQGTGFGFKVIDIQAAPSFIPFDSGGKGTAAVYCLINE
YPREAKVQKVDNALQSGNSQSEVTEQDSKDSYLSLTLLSKADYEKHKVYACEV
T HCLGSLSPVTKSFNRGEC"

```

BASE COUNT	248 a	282 c	229 g	209 t
ORIGIN				
Query Match		89.1%	Score 857;	DB 9; Length 968;
Best Local Similarity		95.4%;	Pred. No. 9,9e-218;	
Matches 894;		Conservative	0; Mismatches	40; Indels
				3; Gaps
				1;

29	GCTCAGTTAGGACCCAGAGGAAACCATGAAAGCCCACTCAGCTTCTCTTCCTCTCTGCT	88
QY		
Db		
13	GCTCAGTTAGGACCCAGAGGAAACCATGAAAGCCCACTCAGCTTCTCTTCCTCTCTGCT	72
QY		
Db		
89	ACTCTGGCTCCACATACCAACCGGAAATTTGTTGACAGACTCTCCAGCACCCCTGTC	148
QY		
Db		
73	ACTCTGGCTCCACGAGTACCAACCGGAAATTTGTTGAGCAGTCTCCAGCACCCCTTTC	132
QY		
Db		
149	TTTGTCTCCAGGGAAAGACCACTCTCTCTGAGGGCCAGTCAAGTGTATACAG- --	205
QY		
Db		
133	TTTGTCTCCAGGGAGAGACCACTCTCTCTCAGGGCCAGTCAAGTGTATGAGCGC	192
QY		
Db		
206	CTACTTAGCTGGTACCAACAGAACTGCGCCAGGCTCCAGGGCTCTCATCTATGATGC	265
QY		
Db		
193	CTACTTAGCTGGTATCAGCAAACTGCGCAGGCTCCAGACTCCTCATGTTTGGTTC	252
QY		
Db		
266	ATCCAACAGGGCACTGGCATCCCAACCAAGTTTCAGTGGAGTGGGTCTGGGACAGACTT	325
QY		
Db		
253	ATCCAGCAGGGCACTGGCATCCCAACAGAGAGTTTCAGTGGAGTGGGTCTGGGACAGACTT	312
QY		
Db		
326	CACCTCTACCATCAGCAGACTGGAGCCCGAAGATGTGGCACTTTATCTACTGTCAACAATA	385
QY		
Db		
313	CACCTCTACCATCAGCAGACTGGAGCCCTGAAGATTTTGCAGTCTATTACTGTCAACAGTA	372
QY		
Db		
386	TTTTTACTACTCCGTACACATTTTGGCCAGGGACAGGCTGGAGATCAACAGCACTGTGCG	445
QY		
Db		
373	TGGTAGTTCACAGGCACATTTCCGGCCCTGGGACCAAGATGGATATCAACAGCAACTGTGCG	432
QY		
Db		
446	TGCACCAATCTGCTTTCATCTCCCGCCATCTGATGAGCAGTTGAAATCTGGAATCCGCTTC	505
QY		
Db		
433	TGCACCAATCTGCTTTCATCTCCCGCCATCTGATGAGCAGTTGAAATCTGGAATCCGCTTC	492
QY		
Db		
506	TGTTGTGTGCTGCTGAATAACTTCTATCCCAAGAGGCGCCAAAGTACAGTGGAGGTGGA	565
QY		
Db		
493	TGTTGTGTGCTGCTGAATAACTTCTATCCCAAGAGGCGCCAAAGTACAGTGGAGGTGGA	552
QY		
Db		
566	TAAAGCCCTCCAAATCGGGTAACTCCAGGAGAGTGTCAAGAGCAGGACAGCAAGGACAG	625
QY		
Db		
553	TAAAGCCCTCCAAATCGGGTAACTCCAGGAGAGTGTCAAGAGCAGGACAGCAAGGACAG	612
QY		
Db		
626	CACCTTACAGCCTCAGCAGCACCTTGACGGTGTGACCAAGCAGACTACAGAGAAACAACAAGT	685
QY		

[illegible]

[illegible]

FEATURES		Location/Qualifiers	
source			
HUMIGGK		1..1244	
LOCUS		/organism="Homo sapiens"	
DEFINITION		/db_xref="taxon:9606"	
BASE COUNT	292 a 383 c 302 g 267 t		
ORIGIN			
Query Match			
Best Local Similarity 82.4%; Score 792.4; DB 6; Length 1244;			
Matches 827; Conservative 0; Mismatches 31; Indels 6; Gaps 1;			
QY	95	GCTCCAGATACACCGGAGAAATTTGTTGACACAGTCTCCAGCCACCTGCTTTGTC	154
Db	381	GATCAACATATACACCGGAGAAATTTGTTGACACAGTCTCCAGCCACCTGCTTTGTC	440
QY	155	TCCAGGGGAAAGACCCACCTCTCTCCAGGCGCCAGTCCAGAGTGTAGCAGTACTTACG	214
Db	441	TCCAGGGGAAAGACCCACCTCTCTCCAGGCGCCAGTCCAGAGTGTAGCAGTACTTACG	500
QY	215	CTGGTACCAACAGAAACCTTCCAGGCTCCAGGCTCTCATCTATGATGATCCCAACAG	274
Db	501	CTGGTACCAACAGAAACCTTCCAGGCTCCAGGCTCTCATCTATGATGATCCCAACAG	560
QY	275	GGCCACTGGCATCCCAACCCAGGTTTCAGTGGCAGTGGGCTTGGGACAGACTTCACTCTAC	334
Db	561	GGCCACTGGCATCCCAACCCAGGTTTCAGTGGCAGTGGGCTTGGGACAGACTTCACTCTAC	620
QY	335	CATCAGCAGCTGGAGCCGCAAGATGTGGCAGTCTTATCTATCTGTCAGC-----AATATT	388
Db	621	CATCAGCAGCTGGAGCCGCAAGATGTGGCAGTCTTATCTATCTGTCAGC-----AATATT	680
QY	389	TACTACTCGGTACATTTTGGCCAGGCGGACAGCTGGAGATCAACAGAACTGTGCTGC	448
Db	681	GCCTCGGCGGCACTTTCCGCGAGGAGCAAGTGGAGATCAACAGAACTGTGCTGC	740
QY	449	ACCATCTGCTTCACTCTCCGCGCATCTGATGAGCAGTTGAATCTGGAAGTCTCTGT	508
Db	741	ACCATCTGCTTCACTCTCCGCGCATCTGATGAGCAGTTGAATCTGGAAGTCTCTGT	800
QY	509	TGTGTGCTGCTGAATAACTTCTATCCAGAGGCGCAAGTACAGTGAAGTGGATAA	568
Db	801	TGTGTGCTGCTGAATAACTTCTATCCAGAGGCGCAAGTACAGTGAAGTGGATAA	860
QY	569	CGCCCTCAATCGGGTAATCCAGGAGAGTGTACAGAGCAGGACCAAGTGGAGTAA	628
Db	861	CGCCCTCAATCGGGTAATCCAGGAGAGTGTACAGAGCAGGACCAAGTGGAGTAA	688
QY	629	CTACAGCCTCAGCAGCAGCTGACCTGAGCAAGACAGTACAGAAACACAAAGTCTA	688
Db	921	CTACAGCCTCAGCAGCAGCTGACCTGAGCAAGACAGTACAGAAACACAAAGTCTA	980
QY	689	CGCCTGGAAGTACCCATCAGGCGCTGAGCTCGGCGGTACAGAGAGCTTCAACAGGG	748
Db	981	CGCCTGGAAGTACCCATCAGGCGCTGAGCTCGGCGGTACAGAGAGCTTCAACAGGG	1040
QY	749	AGAGTGTAGAGGAGAGTGGCCACCTGCTCTCAGTTCAGCCTGACCCCTCCCA	808
Db	1041	AGAGTGTAGAGGAGAGTGGCCACCTGCTCTCAGTTCAGCCTGACCCCTCCCA	1100
QY	809	TCTTTGGCCTCTGACCTTTTTCACAGGGGAGCTTACCCCTATTCGGGTCTCCAGTTC	868
Db	1101	TCTTTGGCCTCTGACCTTTTTCACAGGGGAGCTTACCCCTATTCGGGTCTCCAGTTC	1160
QY	869	ATCTTTTACCTTACCCCTCTCTCTCTGCTTTTAAATGATGATGATGATGATGATGAT	928
Db	1161	ATCTTTTACCTTACCCCTCTCTCTCTGCTTTTAAATGATGATGATGATGATGATGAT	1220
QY	929	TGAATAATAAAGTGAATCTTTTC	952
Db	1221	TGAATAATAAAGTGAATCTTTTC	1244
RESULT 11			

HUMIGGK		1244 bp	mRNA	linear	PRI 27-APR-1993
LOCUS		Human Ig rearranged gamma chain mRNA, V-J-C region and complete cds			
DEFINITION		Human Ig rearranged gamma chain mRNA, V-J-C region and complete cds			
ACCESSION	M63438				
VERSION	M63438.1	GI:184847			
KEYWORDS	C-region; V-region; immunoglobulin gamma-chain; immunoglobulin light chain.				
SOURCE	Human, cDNA to mRNA (transfected into mouse myeloma cell line P3X63Ag8.653).				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Shurford, W., Raff, H.V., Finley, J.W., Esselstyn, J. and Harris, L.J.				
TITLE	Effect of light chain V region duplication on IgG oligomerization and in vivo efficacy				
JOURNAL	Science 252 (5006), 724-727 (1991)				
MEDLINE	91220077				
PUBMED	1902593				
FEATURES					
Source					
	1..1244				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	join(1..49,50..388,389..727,728..1050)				
	/codon_start=1				
	/protein_id="AA73002.1"				
	/db_xref="GI:184847"				
	/translation="MEAPQLFLILLPDTTGEIVLTQSPATLSLSPGERATLSCR				
	ASVSVGLAWYQQKPGQAPRLIVDASNRATGIPARFSGSGGTFTLTISLEPED				
	FVYCYQRDNWPPGATFGGTGKVEIKHTGTEIVLTQSPATLSLSPGERATLSCRASQ				
	SVGSYLAWYQQKPGQAPRLIVDASNRATGIPARFSGSGGTFTLTISLEPEDFAY				
	YCYQRDNWPPGATFGGTGKVEIKHTGTEIVLTQSPATLSLSPGERATLSCRASQ				
	REAKVQWKVDNALASGNSQESVTEQDSKDSYSLSLTLLSKADYEKHKVYACEVTHQ				
	GLSSPYTFSPNRGEC"				
sig_peptide	1..60				
exon	<1..49				
	/note="encodes first part of signal peptide"				
exon	50..388				
	/note="encodes second part of signal peptide and VJ region"				
mat_peptide	61..1047				
exon	389..727				
	/note="duplication of exon 2"				
exon	728..1244				
	/number=3				
	/note="encodes human kappa constant region"				
polya_signal	1223..1232				
BASE COUNT	292 a 383 c 302 g 267 t				
ORIGIN	chromosome 2.				
Query Match					
Best Local Similarity 82.4%; Score 792.4; DB 9; Length 1244;					
Matches 827; Conservative 0; Mismatches 31; Indels 6; Gaps 1;					
QY	95	GCTCCAGATACACCGGAGAAATTTGTTGACACAGTCTCCAGCCACCTGCTTTGTC	154		
Db	381	GATCAACATATACACCGGAGAAATTTGTTGACACAGTCTCCAGCCACCTGCTTTGTC	440		
QY	155	TCCAGGGGAAAGACCCACCTCTCTCCAGGCGCCAGTCCAGAGTGTAGCAGTACTTACG	214		
Db	441	TCCAGGGGAAAGACCCACCTCTCTCCAGGCGCCAGTCCAGAGTGTAGCAGTACTTACG	500		
QY	215	CTGGTACCAACAGAAACCTTCCAGGCTCCAGGCTCTCATCTATGATGATCCCAACAG	274		
Db	501	CTGGTACCAACAGAAACCTTCCAGGCTCCAGGCTCTCATCTATGATGATCCCAACAG	560		
QY	275	GGCCACTGGCATCCCAACCCAGGTTTCAGTGGCAGTGGGCTTGGGACAGACTTCACTCTAC	334		
Db	561	GGCCACTGGCATCCCAACCCAGGTTTCAGTGGCAGTGGGCTTGGGACAGACTTCACTCTAC	620		

Tue Jul 29 09:34:09 2003

```

Db 751 CCACCTGCTCCTCAGCTTCCAGCTGACCCCTCCCATCCTTTGGCCTCTGACCCCTTTTC 810
QY 833 CACAGGGAGCTACCCCTATTTGGGTCCTCCAGCTCATCTTTTCACCTCACCCCTCCTC 892
Db 811 CACAGGGAGCTACCCCTATTTGGGTCCTCCAGCTCATCTTTTCACCTCACCCCTCCTC 870
QY 893 CTCCTTGGCTTTAATTATGCTTAATTTGGGAGGAGATGAATTAATAAGTGAATCTTGC 952
Db 871 CTCCTTGGCTTTAATTATGCTTAATTTGGGAGGAGATGAATTAATAAGTGAATCTTGC 930
QY 953 AAAAAA 962
Db 931 AAAAAA 940

RESULT 14
AR161402
LOCUS 3819 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 393 from patent US 6255458.
ACCESSION AR161402
VERSION AR161402.1 GI:16227274
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3819)
AUTHORS Lonberg, N. and Kay, R.M.
TITLE High affinity human antibodies and human antibodies against digoxin
JOURNAL Patent: US 6255458-A 393 03-JUL-2001;
FEATURES
Location/Qualifiers
1. 3819
/organism="unknown"
BASE COUNT 947 a 1015 c 912 g 945 t
ORIGIN

Query Match 78.18; Score 751; DB 6; Length 3819;
Best Local Similarity 89.5; Pred. No. 2e-189;
Matches 808; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 51 ACCATGGAAGCCCGCCAGCTCAGCTTCTTCTCCTCCTGCTACTTGGCTCCAGATACACC 110
Db 2443 ACCATGATGTTCCAGCTCAGCTCCTCGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2502
QY 111 GGAGAAATGTTGTCACAGCTCTCCAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 170
Db 2503 TCGCATCATCCAGATGACCCAGCTCTCCATCTTCGCTGCTGCTGCTGCTGCTGCTGCTG 2562
QY 171 ACCCTCTCTCGAGGCGCCAGTCAGAGTGTAGGAGCTACTTACGCTGGTACCAACAGAAA 230
Db 2563 ACCATCATTTGTGCGGCGAGTCAGGATATTAGCAGCTGGTGTAGCTGCTGCTGCTGCTG 2622
QY 231 CTTGGCCAGGCTCCAGGCTCCTCATCTATGATGATCCACAGAGGCGCACTGGCATCCCA 290
Db 2623 CCAGGTAAGACCTTAAGCTCTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATG 2682
QY 291 CCCAGGTTCAAGTGGGAGTGGGTCCTGGGAGAGCTTCACTCTCAGCATCAGAGCTGGAG 350
Db 2683 TCAGGTTCAAGGGAAGTGGATCTGGGAGAGATTTCACTCTCAGCATCAGAGCTGGAG 2742
QY 351 CCCAGAGATGTCGCACTTTATTTACTCTCAGCAATATTTTACTTACTTCGCTACATTTTGC 410
Db 2743 CCTGAAGATTTGCACTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2802
QY 411 CAGGGGACAGGCTGGAGATCAACGAATGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 470
Db 2803 CAGGGAACCAAGCTGGAGATCAACGAATGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2862
QY 471 CCACTGTATGAGAGTTGAATCTGGAATCTGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 530
Db 2863 CCATCTGATGAGCAGTTGAATCTGGAATCTGGAATCTGCTGCTGCTGCTGCTGCTGCTG 2922
QY 531 TATCCAGAGAGGCCAAAGTACAGTGAAGGTGATTAAGCCCTCCCAATCGGGTAACCTCC 590

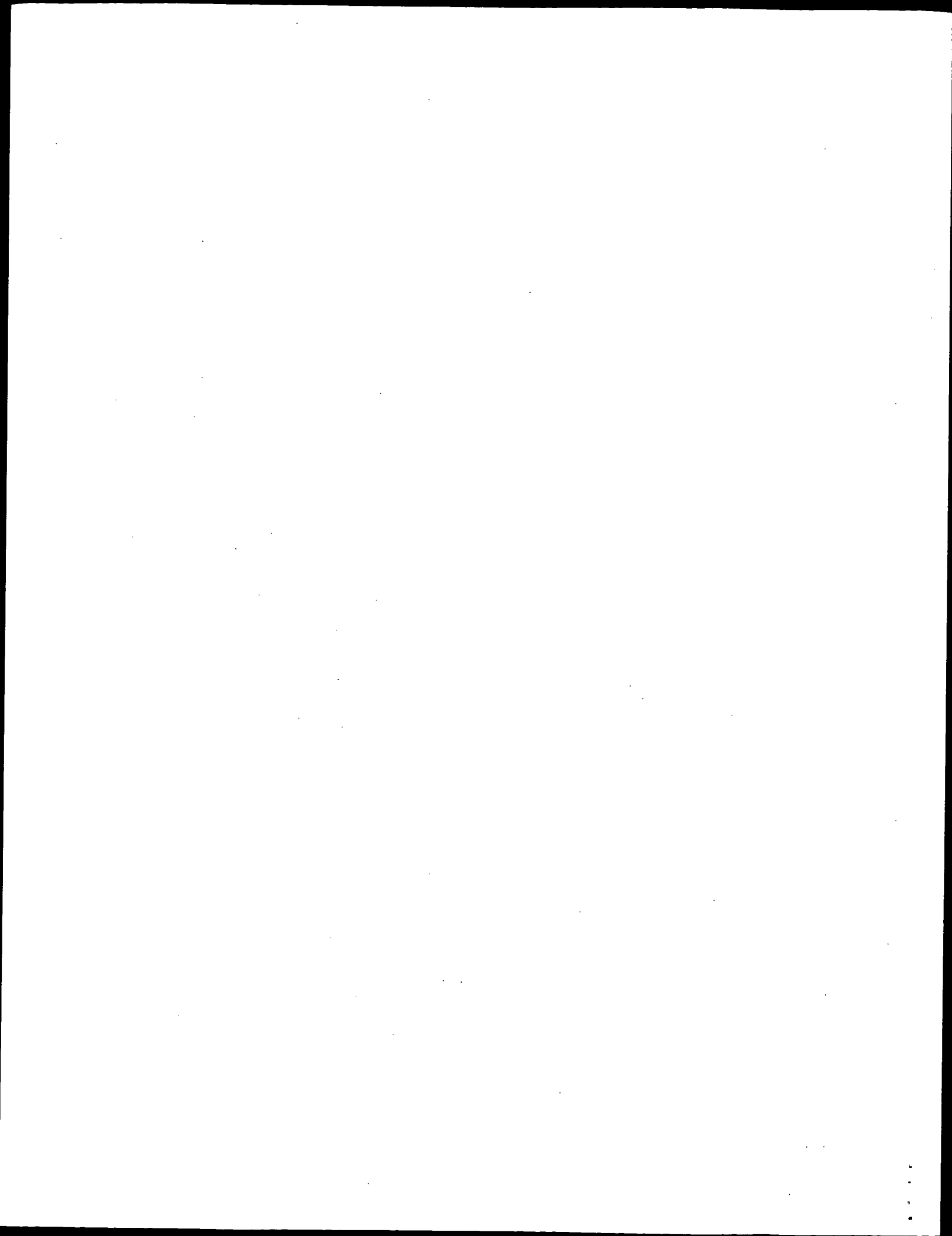
```

```

Db 2923 TATCCAGAGAGGCCAAAGTACAGTGAAGGTGGATATACGCCCTCCAAATCGGGTAACCTCC 2982
QY 591 CAGGAGAGTGTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 650
Db 2983 CAGGAGAGTGTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3042
QY 651 AGCCTTGACAAAGAGAGAGTACGAGAAACACAAAGTCTTACGCTTGGGCTGCGAGTACCCATCAG 710
Db 3043 AGCCTTGAGAAAGAGAGTACGAGAAACACAAAGTCTTACGCTTGGGCTGCGAGTACCCATCAG 3102
QY 711 GGCCTGAGCTGCCCGTCCACAAAGAGTTCACAGGGGAGAGTGTAGAGGGGAGAGTGC 770
Db 3103 GGCCTGAGCTGCCCGTCCACAAAGAGTTCACAGGGGAGAGTGTAGAGGGGAGAGTGC 3162
QY 771 CCCACCTGCTCCTCAGTTCAGCTTCCAGCTGACCCCTCCCATCCTTTGGGCTCTGACCCCTTT 830
Db 3163 CCCACCTGCTCCTCAGTTCAGCTTCCAGCTGACCCCTCCCATCCTTTGGGCTCTGACCCCTTT 3222
QY 831 TCCACAGGGAGCTACCCCTATTTGGGCTTCCAGCTCATCTTTTCACCTCACCCCTCC 890
Db 3223 TCCACAGGGAGCTACCCCTATTTGGGCTTCCAGCTCATCTTTTCACCTCACCCCTCC 3282
QY 891 TCCTCTGCTTGTATTAATTATGCTTAATTTGGGAGGAGATGAATTAATAAGTGAATCTTT 950
Db 3283 TCCTCTGCTTGTATTAATTATGCTTAATTTGGGAGGAGATGAATTAATAAGTGAATCTTT 3342
QY 951 GCA 953
Db 3343 GCA 3345

RESULT 15
BC005332 953 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, Similar to immunoglobulin kappa constant, clone
DEFINITION MGC:12418 IMAGE:3934658, mRNA, complete cds.
ACCESSION BC005332
VERSION BC005332.1 GI:13529115
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 953)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL plate: 16 Row: d Column: 21.
Location/Qualifiers
1. .953
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:12418 IMAGE:3934658"
/tissue_type="Skeletal Muscle"
/clone_lib="NIH_MGC_81"
FEATURES
source

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 13:53:40 ; Search time 221.852 Seconds
(without alignments)
9765.155 Million cell updates/sec

Title: US-09-806-276A-3
Perfect score: 962
Sequence: 1 tcgagcgcattcgctcgag.....gaattcttgcacaaaaaaa 962

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	962	100.0	962	21	AAA09154 Human BMDSP-1 codi
2	880.4	91.5	1061	22	AAH31379 Human secreted pro
3	845.4	87.9	970	24	AAS99477 Anti-human AILIM m
4	839.6	87.3	924	15	AAQ71872 Sequence coding hu
5	830.2	86.3	948	24	AAS99475 Anti-human AILIM m
6	829.6	86.2	931	22	AAC66524 Human immune syste
7	808.4	84.0	913	24	AAS62744 CDNA sequence #531
8	792.8	82.4	1450	22	AAS22766 Human cDNA encodin
9	792.8	82.4	1450	22	AAS22767 Human cDNA encodin

10	792.8	82.4	1450	22	AAS22768 Human cDNA encodin
11	792.8	82.4	1450	22	AAS22769 Human cDNA encodin
12	792.4	82.4	1242	12	AQ11878 lpl IgG aberrant l
13	792.4	82.4	1244	24	ABK64543 Human benign prost
14	792.4	82.4	1244	24	ABL65461 Lung cancer relate
15	771.2	80.2	1721	22	AAS22530 Human cDNA encodin
16	771.2	80.2	1736	22	AAS22532 Human cDNA encodin
17	771.2	80.2	1741	22	AAS22531 Human cDNA encodin
18	761.6	79.2	1106	24	ABQ54241 Human ovarian anti
19	751	78.1	3819	18	AA778825 Kappa light chain
20	751	78.1	3819	19	AAV39266 Plasmid pLC6G5 nuc
21	751	78.1	3819	20	AAZ22020 Nucleotide sequenc
22	748	77.8	936	21	AAZ27390 Human IGFAM-10 imm
23	746.8	77.6	974	24	AAS99473 Anti-human AILIM m
24	738.8	76.8	1033	22	AAS00145 Human cDNA clone H
25	738.8	76.8	1066	14	AAQ49943 Human anti-HBs lig
26	735.6	76.5	944	22	AA44892 Human breast cance
27	727	75.6	917	21	AAZ27381 Human IGFAM-1 immu
28	725.8	75.4	1517	22	AAH98191 Human EST-derived
29	724.6	75.3	990	22	AAF97952 Human secreted pro
30	723.6	75.2	930	22	AAH31350 Human secreted pro
31	721.4	75.0	928	21	AAZ27393 Human IGFAM-13 imm
32	718.4	74.7	1404	22	AAS22861 Human cDNA encodin
33	717.6	74.6	923	21	AA95785 Human immune syste
34	713.2	74.1	941	22	AAF97954 Human secreted pro
35	710.4	73.8	928	22	AAH26799 Human immune respo
36	709.4	73.8	916	21	AAZ27385 Human IGFAM-5 immu
37	705	73.3	851	22	AAH98251 Human EST-derived
38	685	71.2	958	22	AA66520 Human immune syste
39	683.6	71.1	802	22	AA44872 Human breast cance
40	680	70.7	945	13	AAQ20067 Encodes light chal
41	680	70.7	963	22	AAF44895 Human breast cance
42	662.2	68.8	957	15	AAQ54654 chiT84.12 L6 light
43	658.2	68.4	890	21	AAZ27392 Human IGFAM-12 imm
44	656.8	68.3	708	24	AAK98702 CDNA of the light
45	652.2	67.8	705	21	AAA46871 DNA encoding the k

ALIGNMENTS

RESULT 1
AAA09154
ID AAA09154 standard; DNA; 962 BP.

XX AAA09154;

AC 10-AUG-2000 (first entry)

XX Human BMDSP-1 coding sequence.

DE Bone marrow-derived serum protein; immunoglobulin kappa light chain;

XX BMDSP-1; cytostatic; anti-HIV; antiatherosclerotic;

KW anti-inflammatory; antiarthritic; antisclerotic; hypotensive; antiviral;

KW antiparasitic; ss...

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 54..758

FT /*tag= a

FT /product= BMDSP-1

XX WO200020588-A2.

XX 13-APR-2000.

XX 01-OCT-1999; 99WO-US22908.

XX 02-OCT-1998; 98US-0165621.

XX (INCY-) INCYTE PHARM INC.

Yoda

Query Match	100.0%	Score 962;	DB 21;	Length 962;
Best Local Similarity	100.0%;	Pred. No. 1.7e-240;		
Matches 962;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TCGAGCCGATTCGGCTCGAGCGCTCGAGCTCAGTTAGGACCCAGAGGGAACCATGGAG	60	
DB	1			
QY	1	TCGAGCCGATTCGGCTCGAGCGCTCGAGCTCAGTTAGGACCCAGAGGGAACCATGGAG	60	
DB	1			
QY	61	CCCAGCTCAGCTTCCTCTCCTGCTACTCTCTGGCTCCAGATACCCACCGGAGAAATTG	120	
DB	61	CCCAGCTCAGCTTCCTCTCCTGCTACTCTGGCTCCAGATACCCACCGGAGAAATTG	120	
QY	121	TGTTGACACAGTCTCCAGCCACCTGCTCTTGTCTCCAGGGAAAGAGCCACCTCTCCT	180	
DB	121	TGTTGACACAGTCTCCAGCCACCTGCTCTTGTCTCCAGGGAAAGAGCCACCTCTCCT	180	
QY	181	GCAGGGCCAGTCAGAGTGTGTAGCAGTACTTAGCTGTGTACCAACAGAAACTGSGCCAGG	240	
DB	181	GCAGGGCCAGTCAGAGTGTGTAGCAGTACTTAGCTGTGTACCAACAGAAACTGSGCCAGG	240	
QY	241	CTCCAGGCTTCCTCATCTATGATGATCAACAGGGCCACTGGCATCCACCCAGGTTCA	300	
DB	241	CTCCAGGCTTCCTCATCTATGATGATCAACAGGGCCACTGGCATCCACCCAGGTTCA	300	
QY	301	GTGGCAGTGGGTCTGGGACAGACTTCACCTCTCACCATCAGCAGACTGGAGCCGCAAGATG	360	
DB	301	GTGGCAGTGGGTCTGGGACAGACTTCACCTCTCACCATCAGCAGACTGGAGCCGCAAGATG	360	
QY	361	TGGCACTTTATTACTGTACGAATATTTTACTACTCCGTACACTTTTGGCCAGGGGACCA	420	
DB	361	TGGCACTTTATTACTGTACGAATATTTTACTACTCCGTACACTTTTGGCCAGGGGACCA	420	
QY	421	GGCTGGAGATCAAGCAACTGTGGCTGCACCATCTGCTTTCATCTTCCCGCATCTGATG	480	
DB	421	GGCTGGAGATCAAGCAACTGTGGCTGCACCATCTGCTTTCATCTTCCCGCATCTGATG	480	
QY	481	AGCAGTTGAATCTCGAACTGCCTCTGTTGTGGCTGCTGAATAACTTTCATCCCAAG	540	
DB	481	AGCAGTTGAATCTCGAACTGCCTCTGTTGTGGCTGCTGAATAACTTTCATCCCAAG	540	
QY	541	AGGCCAAAGTACAGTGGAGGTGGATAACGGCCCTCCAATCGGGTAACTCCCAAGSAGATG	600	
DB	541	AGGCCAAAGTACAGTGGAGGTGGATAACGGCCCTCCAATCGGGTAACTCCCAAGSAGATG	600	
QY	601	TACAGCAGCAGCAGCAGGAGCAGACACCTACAGCCTCAGCAGCACCTGACGCTGAGCA	660	
DB	601	TACAGCAGCAGCAGCAGGAGCAGACCTACAGCCTCAGCAGCACCTGACGCTGAGCA	660	

Human secreted protein-encoding gene 2 CDNA clone HTSER67, SEQ ID NO:41.

WO200132674-A1.

10-MAY-2001.

25-OCT-2000: 2000W0-11529360

20-00000-1000.

30-JUN-2000; 2000US-0215138.

(HUMA -) HUMAN GENOME SCI INC

Ruben SM, Komatsu J, Young DE, Moore DA.

WPT: 2001-201051/30

N-PSDB; AAG71272.

New nucleic acid molecule encoding a human secreted protein, useful for preventing, treating or ameliorating medical conditions such as rheumatoid arthritis, Alzheimer's disease and microbial infections -

Claim 4; Page 447; 581pp; English.

AAH31349-AAH31428 represent cDNAs corresponding to 26 human secreted

Db 601 CCTGACGCTGAGCAAGACAGACTACAGAGAAACACAAAGTCTACGCTCGAAGTCAACCA 660
 QY 707 TCAGGGCCTGAGCTCGCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTAGAGGAGAA 766
 Db 661 TCAGGGCCTGAGCTCGCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTAGAGGAGAA 720
 QY 767 GTGCCCCCAGCTGCTCCTCAGTTCAGCTGACGCTGACCCCTCCCATCCTTTGGGCTCTGACCC 826
 Db 721 GTGCCCCCAGCTGCTCCTCAGTTCAGCTGACGCTGACCCCTCCCATCCTTTGGGCTCTGACCC 780
 QY 827 TTTTTCACAGGGGACCTACCCCTATGCGGTCTCCAGCTCATCTTTCACTCACCC 886
 Db 781 TTTTTCACAGGGGACCTACCCCTATGCGGTCTCCAGCTCATCTTTCACTCACCC 840
 QY 887 CTCCTCCTCTTGGCTTTAATTATGCTAATGTTGGAGGAGATGAATAAATAAGTGAAT 946
 Db 841 CTCCTCCTCTTGGCTTTAATTATGCTAATGTTGGAGGAGATGAATAAATAAGTGAAT 900
 QY 947 CTTTGC 952
 Db 901 CTTTGC 906

RESULT 8

AA522766
 ID AAS22766 standard; cDNA; 1450 BP.

XX AAS22766;

DT 24-OCT-2001 (first entry)

DE Human cDNA encoding a novel human protein #332.

XX Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;
 KW immunomodulatory; cytostatic; neuroprotective; vulnery; nootropic;
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KW tissue regeneration; immune disorder.

XX Homo sapiens.

XX WO200155437-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02623.

XX 25-JAN-2000; 2000US-0491404.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-451939/48.

XX P-PSDB; AAU14461.

PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -

XX Claim 1; Page 716; 894pp; English.

XX The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or

CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/elicit an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence encodes a protein of the invention.

XX Sequence 1450 BP; 332 A; 401 C; 395 G; 322 T; 0 other;

Query Match 82.4%; Score 792.8; DB 22; Length 1450;
 Best Local Similarity 95.9%; Pred. No. 2.1e-196;
 Matches 825; Conservative 0; Mismatches 32; Indels 3; Gaps 1;

QY 106 CCACCGGAGAAATGTGTTGACACAGCTCCAGCCACCCTGTCTTGTCTCCAGGGGAAA 165
 Db 591 CGACTGGCGAAATGTGTTGACGCGAGTCTCCAGGCACTCTTTGTCTCCAGGGGAAA 650
 QY 166 GAGCCACCTCTCTCGCAGGGCCAGTCAGAGTGT---AGCAGCTACTTAGCTGTACC 222
 Db 651 GAGCCACCTCTCTCGCAGGGCCAGTCAGAGTGT---AGCAGCTACTTAGCTGTACC 710
 QY 223 AACAGAAACCTGGCCAGGCTCCAGGCTCCTCATCTATGATGCATCAACAGGGCCACTG 282
 Db 711 AGCAGAAACCTGGCCAGGCTCCAGGCTCCTCATCTATGATGCATCAACAGGGCCACTG 770
 QY 283 GCATCCACCCAGGTTCAAGTGGAGTGGTCTGGGACAGACTTCATCTCACCATCAGCA 342
 Db 771 GCATCCACCCAGGTTCAAGTGGAGTGGTCTGGGACAGACTTCATCTCACCATCAGCA 830
 QY 343 GACTGGAGCCGCAAGATGTGGCACTTATTTACTGTGTCAGCAATATTTTACTCTCCGTACA 402
 Db 831 GACTGGAGCCGCAAGATGTGGCACTTATTTACTGTGTCAGCAATATTTTACTCTCCGTACA 890
 QY 403 CTTTGGCCAGGGGACCCAGGCTGGAGTCAACAGAACTGTGGCTGCACCATCTGTCTTCA 462
 Db 891 CGTTCGGCCAAAGGACCAAGTGGATCAAAAGAACTGTGGCTGCACCATCTGTCTTCA 950
 QY 463 TCTTCCCGCATCTGATGAGCAGTTGAAATCTGGAATCTGCTCTGTGTGTGCTGCTGA 522
 Db 951 TCTTCCCGCATCTGATGAGCAGTTGAAATCTGGAATCTGCTCTGTGTGTGCTGCTGA 1010
 QY 523 ATAACTTCTATCCAGAGAGGCCAAAGTACAGTGGAGGTGGATACAGCCCTCCAATCGG 582
 Db 1011 ATAACTTCTATCCAGAGAGGCCAAAGTACAGTGGAGGTGGATACAGCCCTCCAATCGG 1070
 QY 583 GTAACCTCCAGGAGAGTGTACAGAGCAGACAGCAAGGACAGCACCCTACAGCCTCAGCA 642
 Db 1071 GTAACCTCCAGGAGAGTGTACAGAGCAGACAGCAAGGACAGCACCCTACAGCCTCAGCA 1130
 QY 643 GCACCTCGAGCTGACAAAGAGACTACAGAGAAACACAAAGTCTACGCTCTCGAAGTCA 702
 Db 1131 GCACCTCGAGCTGACAAAGAGACTACAGAGAAACACAAAGTCTACGCTCTCGAAGTCA 1190
 QY 703 CCATCAGGGCCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTAGAGGG 762
 Db 1191 CCATCAGGGCCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTAGAGGG 1250
 QY 763 AGAAGTGGCCCACTGCTCTCAGTTCAGGCTGACCCCTCCCATCTCTTTGGGCTCTG 822
 Db 1251 AGAAGTGGCCCACTGCTCTCAGTTCAGGCTGACCCCTCCCATCTCTTTGGGCTCTG 1310
 QY 823 ACCCTTTTTCACAGGGGACCTACCCCTATTGCGGTCTCTCCAGCTCATCTTTTCACTCAC 882
 Db 1311 ACCCTTTTTCACAGGGGACCTACCCCTATTGCGGTCTCTCCAGCTCATCTTTTCACTCAC 1370

CC anti-inflammatory diseases, nervous system disorders, and infection.
CC The present sequence encodes a protein of the invention.
XX
S0 Sequence 1450 BP; 332 A; 401 C; 395 G; 322 T; 0 other;

PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX (AVAL-) AVALON PHARM.
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 DR Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX Claim 1; SEQ ID 3798; 44pp; English.
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX Sequence 1244 BP; 292 A; 383 C; 302 G; 267 T; 0 other;
 SQ

Query Match 82.4%; Score 792.4; DB 24; Length 1244;
 Best Local Similarity 95.7%; Pred. No. 2.6e-196;
 Matches 827; Conservative 0; Mismatches 31; Indels 6; Gaps 1;

QY 95 GCTCCAGATACACCGGAAATTTGTTGACAGCTCCAGGACACCTGCTCTTTGTC 154
 DB 381 GATCAAAATACACCGGAAATTTGTTGACAGCTCCAGGACACCTGCTCTTTGTC 440
 QY 155 TCAGGGGAAAGAGCCACCTCTCTGCGAGGCGGAGTGTAGAGCTACTTACG 214
 DB 441 TCAGGGGAAAGAGCCACCTCTCTGCGAGGCGGAGTGTAGAGCTACTTACG 500
 QY 215 CTGGTACCACAGAAACCTGGCCAGGCTCCAGGCTCTCATCTATGATGATCAACAG 274
 DB 501 CTGGTACCACAGAAACCTGGCCAGGCTCCAGGCTCTCATCTATGATGATCAACAG 560
 QY 275 GGCCACTGCATCCACCCAGGTTTCAGTGGCAGTGGGCTGGGACAGACTTCACTCTAC 334
 DB 561 GGCCACTGCATCCACCCAGGTTTCAGTGGCAGTGGGCTGGGACAGACTTCACTCTAC 620
 QY 335 CATCAGCAGCTGGAGCCGAGATGTGCGACTTTTACTGTCAGC-----AATATT 388
 DB 621 CATCAGCAGCTGGAGCTGAGATTTTTCAGTGTGATGATGATGATGATGATGATG 680
 QY 389 TACTACTCCGTACACTTTTGGCCAGGAGCCAGGCTGGAGATCAAGAACTGTGGCTGC 448
 DB 681 GCCTCCGGGGCCACTTTTCGGGGAGGAGCAAGGTGGAGATCAAGAACTGTGGCTGC 740
 QY 449 ACCATCTGCTTCATCTTCCCGGCTCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGT 508
 DB 741 ACCATCTGCTTCATCTTCCCGGCTCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGT 800
 QY 509 TGTGTGCTGCTGAATACTTCTATCCAGAGAGGCCAAAGTACAGTGGAGGTGATAA 568
 DB 801 TGTGTGCTGCTGAATACTTCTATCCAGAGAGGCCAAAGTACAGTGGAGGTGATAA 860
 QY 569 CGCCCTCCAAATCGGGTAACTCCAGGAGAGTGTACAGAGCAGCAGCAAGGACAGCAC 628

DB 861 CGCCCTCCAAATCGGTAACCTCCAGGAGAGTGTACAGAGCAGGACAGCAAGGACAGCAC 920
 QY 629 CTACAGCCTCAGCAGCACCCTGAGCTGAGCAAAAGCAGACTACGAGAAACACAAAGTCTTA 688
 DB 921 CTACAGCCTCAGCAGCACCCTGAGCTGAGCAAAAGCAGACTACGAGAAACACAAAGTCTTA 980
 QY 689 CGCCTCGGAAGTCAACCCATCAGGGCTGAGCTCGCCCTGACAAAAGAGCTTCAACAGGGG 748
 DB 981 CGCCTCGGAAGTCAACCCATCAGGGCTGAGCTCGCCCTGACAAAAGAGCTTCAACAGGGG 1040
 QY 749 AGAGTGTAGAGGAGAGAGTGGCCCACTGCTCTCAGTTCAGCTGACCCCTCCCA 808
 DB 1041 AGAGTGTAGAGGAGAGAGTGGCCCACTGCTCTCAGTTCAGCTGACCCCTCCCA 1100
 QY 809 TCCTTTGGCCTCTGACCCCTTTTTCACAGGGGACCTACCCCTATTTCGCGTCTCCAGCTC 868
 DB 1101 TCCTTTGGCCTCTGACCCCTTTTTCACAGGGGACCTACCCCTATTTCGCGTCTCCAGCTC 1160
 QY 869 ATCTTTCACTCACCCT 928
 DB 1161 ATCTTTCACTCACCCT 1220
 QY 929 TGAATAAATAAAGTGAATCTTTGC 952
 DB 1221 TGAATAAATAAAGTGAATCTTTGC 1244

RESULT 15
 AAS22530
 ID AAS22530 standard; cDNA; 1721 BP.
 XX AC AAS22530;
 XX DT 24-OCT-2001 (first entry)
 XX DE Human cDNA encoding a novel human protein #96.
 KW Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;
 KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;
 KW anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KW tissue regeneration; immune disorder.
 OS Homo sapiens.
 XX WO200155437-A2.
 XX 02-AUG-2001.
 XX 25-JAN-2001; 2001WO-US02623.
 XX 25-JAN-2000; 2000US-0491404.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 DR WPI; 2001-451939/48.
 DR P-PSDB; AAU14225.
 XX Isolated polypeptides useful for treating anti-inflammatory diseases,
 XX nervous system disorders, and for regenerating bone and cartilage -
 PS Claim 1; Page 301-302; 894pp; English.
 CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as

CC molecular weight markers, food supplements, and in antibody production.
CC The polypeptides are used to identify compounds which bind to the
CC polypeptides. Polynucleotides of the invention are used as probes and
CC primers, for sequencing, for chromosome or gene mapping, in the
CC production of recombinant proteins, and in generating anti-sense DNA or
CC RNA and in gene therapy. Polypeptides of the invention can be used to
CC target drugs to a tumour. In assays to determine biological activity, to
CC raise antibodies/elicit an immune response, to determine quantitative
CC protein levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue; wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
CC The present sequence encodes a protein of the invention.
XX
SQ

Sequence 1721 BP; 394 A; 473 C; 428 G; 426 T; 0 other;
Query Match 80.2%; Score 771.2; DB 22; Length 1721;
Best Local Similarity 95.7%; Pred. No. 9.4e-191; Indels 4; Gaps 2;
Matches 815; Conservative 0; Mismatches 33;
QY 106 CCACCGGAGAAATGTTGACACAGTCTCCAGCACCCTGTCTTCTCCAGGGGAAA 165
DB 601 CGACTGGCGAAATGTTGACGAGTCTCCAGGACCCCTGTCTTCTCCAGGGGAAA 660
QY 166 GAGCCACCCCTCTCTCGAGGGCCAGTCAGAGTGTT--AGCAGTACTAGCCCTGGTACC 222
DB 661 GAGCCACCCCTCTCTCGAGGGCCAGTCAGAGTGTTAGCAGCAGTACTTAGCCCTGGTACC 720
QY 223 AACAGAACTGGCCAGGCTCCAGGCTCCCTCATCTATGATGATGATCCAAACAGGGCCACTG 282
DB 721 AGCAGAACTGGCCAGGCTCCAGGCTCCAGGCTCCCTCATCTATGATGATGATCCAAACAG 780
QY 283 GCATCCACCCAGGTTTCAGTGGCAGTGGTCTGGGACAGACTTCACTCTCACCATCAGCA 342
DB 781 GCATCCACAGAGTTTCAGTGGCAGTGGTCTGGGACAGACTTCACTCTCACCATCAGCA 840
QY 343 GACTGGAGCCGGAAGATGAGGACATTTACTGTGACCAATATTTTACTACTCGGTACA 402
DB 841 GACTGGAGCCGGAAGATTTGACAGTATTTACTGTGACAGAGGCTGCTATTCGCGCA 900
QY 403 CTTTGGCCAGGGGACCGAGCTGGAGATCAACAGAACTGTGGCTGCACCACTGTCTTCA 462
DB 901 CGTTGGCCAGGAGGACCAAGGTGGAATCAACAGAACTGTGGCTGCACCACTGTCTTCA 960
QY 463 TCCTCCCGCATCTGATGACAGCTTGAATCTGGAATCTGCTCTGTGTGCTGCTGCTGA 522
DB 961 TCCTCCCGCATCTGATGACAGTGAATCTGGAATCTGCTCTGTGTGCTGCTGCTGA 1020
QY 523 ATAACTTCTATCCAGAGGCGCAAGTACAGTGGAGGTGGATTAACGCCCT-CCAAATCG 581
DB 1021 ATAACTTCTATCCAGAGGCGCAAGTACAGTGGAGGTGGATTAACGCCCTCCCAATCG 1080
QY 582 GGTAACTCCAGAGAGTGTTCACAGCAGGAGGACAGCAGCAGCAGCAGCAGCAGCAGCAGC 641
DB 1081 GGTAACTCCAGAGAGTGTTCACAGCAGGAGGACAGCAGCAGCAGCAGCAGCAGCAGCAG 1140
QY 642 AGCACCCTGAGCTGAGCAAAAGCAGTACGAGAAACACAAAGTCTACGCTCGGAAGTC 701
DB 1141 AGCACCCTGAGCTGAGCAAAAGCAGTACGAGAAACACAAAGTCTACGCTCGGAAGTC 1200
QY 702 ACCCATCAGGGCTGAGCTCGCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTAGAGG 761
DB 1201 ACCCATCAGGGCTGAGCTCGCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTAGAGG 1260
QY 762 GAGAAGTGGCCCACTCTCTCAGTTCCAGCTGACCCCTCCCATCTCTTGGGCTCT 821

DB 1261 GAGAAGTGGCCCACTCTCTCAGTTCCAGCTGACCCCTCCATCTCTTGGGCTCT 1320
QY 822 GACCCTTTTCCACAGGGGACCTACCCCTATTGCGGTCTCTCCAGCTCATCTTTCACCTCA 881
DB 1321 GACCCTTTTCCACAGGGGACCTACCCCTATTGCGGTCTCTCCAGCTCATCTTTCACCTCA 1380
QY 882 CCCCCCTCTCTCTCTTGGCTTTTAAATATGCTAATGTTGGAGGAGAATGAATAAAG 941
DB 1381 CCCCCCTCTCTCTCTTGGCTTTTAAATATGCTAATGTTGGAGGAGAATGAATAAAG 1440
QY 942 TGAATCTTTGCA 953
DB 1441 TGAATCTTTGCA 1452

Search completed: July 25, 2003, 15:52:49
Job time : 227.852 secs

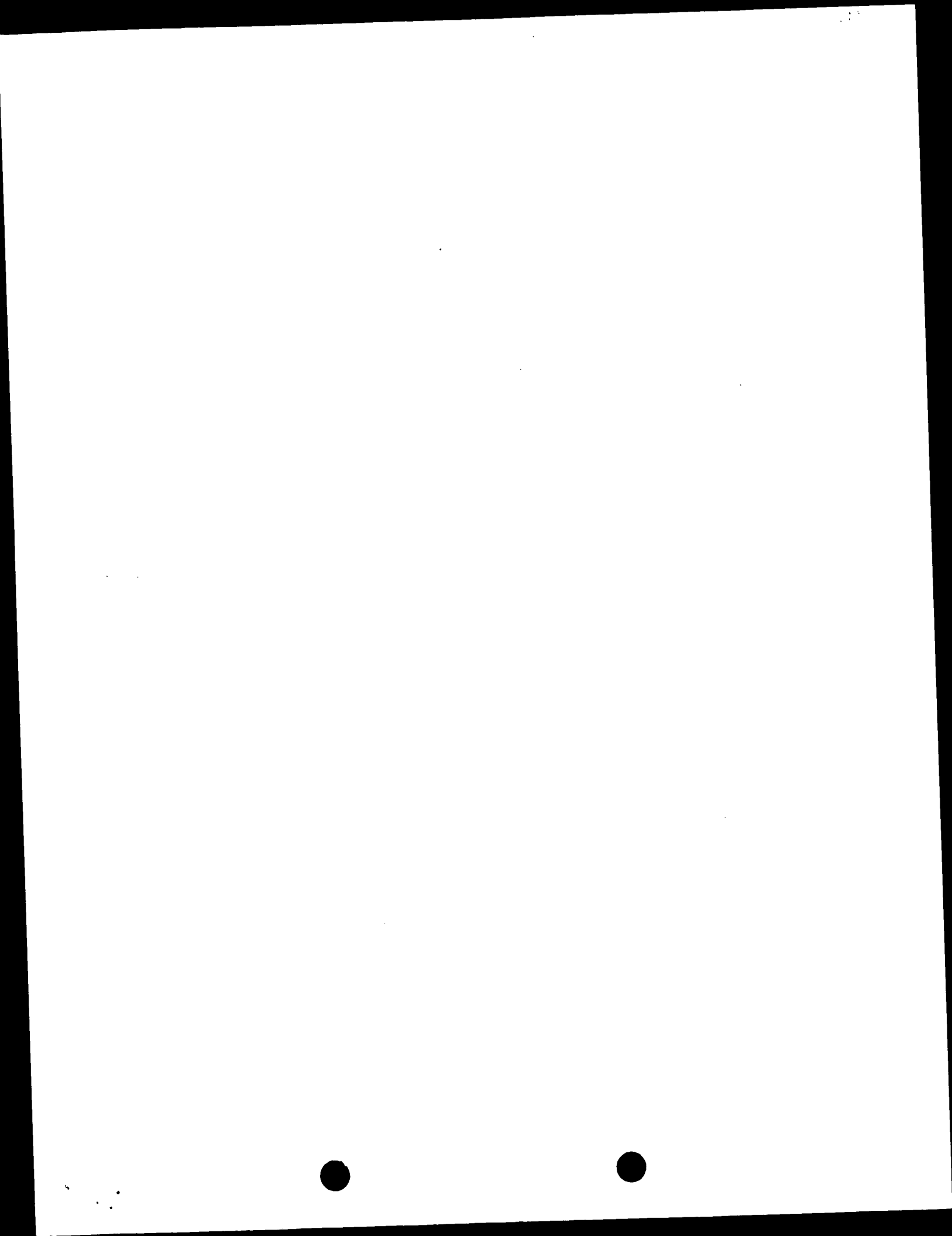
Query Match 30.2%; Score 769.8; DB 10; Length 1481;
Best Local Similarity 75.2%; Pred. No. 7.6e-177;
Matches 1062; Conservative 0; Mismatches 297; Indels 54; Gaps 6;

QY 20 CGGCGCCGCGGCGCTTCCAGCGCGCTCTCGGAGCGGATATCGCGGTGAGCAGGCA 79
DB 60 CGGCTGCTCCGCGGAGCAGCGCGCGCGAGTCCCGGAGCAGATCTGCGGTGAACCGTTG 119
QY 80 CCAGCGCGGAGACTCGGCTGGATCTGCG-----CACAGCGGCGGAGGATTCGCT 128
DB 120 CGGCGCTCTGCTGCTCGCGCAGCAGCGAGGACTGTTGGCCCTCGGGAGCGCACCGG 179
QY 129 GCGCGCGCGGAGCGCGCGCGCGAGCGCTCGGATCTCAGCGCGCGCGGTTGCTCTGG 188
DB 180 GCGAGCGCGCGCGCTCCGCGGAGAGAGCTAGGAGCGCTTAGAGACTGCTGCTTCTCCGAG 239
QY 189 TTGTGCTCAAGACTGGATGATGTAAGTGGCTCTTAGGAGCGCTCACTTGGCGGTAACT 248
DB 240 CTGTGGAAGGACTCGATATATGTAAGTGGCTCTCCGGAAGAGCGCGCGCTCGCT 299
QY 249 CAGGAAGGTTCTCTTTGACCCCACTATCTGGAAGCCACTTCTGAGCCACTTGAAGAA 308
DB 300 GAGCAAGATCTCCAGACCGCTGCTGCTTTTAGAGCACTTGAGAAG-----A 348
QY 309 AATGATGTGACAGTCTATCAAAAAGGA---TTCAGAAACATATACCATCTGTGAAGAA 365
DB 349 AATGATGTGACAAATCTTATGAAAAAGGAGCTTTTGGAAAGTGTAGGATACCGCTGCAGAA 408
QY 366 AGTGGCGCTTCTCCGCGCTTGCAGAAATAGACATTTCAAAATCCAAATGCCAGCCAGA 425
DB 409 GGTCACTCTTCCCTGTGTGTAATAGACATCTCAATTCCTCAAAATGCCGTGCCAGA 468
QY 426 CCCAATTTACCTGAAAGCAGCAATPACAAGAAAGAAAGAAATTTAACTAGAGGACA 485
DB 469 CCCCAATTTACCTGAAAGCAGCTTAACAAGAAAGAAAGAAATTTAACTAGAGGACA 528
QY 486 TTCTGTCTCTGATATGATCAGTCCCGCTTGGAGACTTTGCCACACCATCCACATTG 545
DB 529 TTCTGTCTCTGATATGATCAGTCCCGCTTGGAGACTTTGCCACACCATCCACATTG 588
QY 546 GCAAGAGGCGCCAGCAGTGTCTTTGGAGATATTTCTTCTCAAGGAACTACGAGC 605
DB 589 GCAAGAGGCGCCAGCAGTGTCTTTGGAGATATTTCTTCTCAAGGAACTATGAGC 648
QY 606 TTTTACTCTGGAACAGCAGGAGAACACACTCGGCGAGTTCCCTGGGCATATAGTTCT 665
DB 649 TCTTGGCGGAAACCAAGAGAGGCACACTCTGGCGAGTTCCCGGGCATACGACTTCT 708
QY 666 TCGGCGCCCAACAGCAGCTCGGACTCTGTGTTACAGAAAGCGCTCCCGGTGCTCAAAA 725
DB 709 TCGGCGCCCAACAGCAGCTCGGACTCTGTGTTACAGAAAGCGCTCCCGGTGCTCAAAA 768
QY 726 ATGCCATCTCTCCGACCATTTGGAGGATCCCAAGCTCTCATGTTGGCCCTTATTTGTCAC 785
DB 769 ATGCCATCTCTCCGACCATTTGGAGGATCCCAAGCTCTCATGTTGGCCCTTATTTGTCAC 828
QY 786 CAGTGACATTTATTTCAACAGAGGAGTCTTCCGCGCAGCAAGCTGCCAGGCTTAGCT 845
DB 829 CAGTGACATTTATTTCAACAGAGGAGTCTTCCGCGCAGCAAGCTGCCAGGCTTAGCT 888
QY 846 GCGAGCGCGCTCATGAGGAGAAAGCTCAGGAGAAAGCAGTCTGTGAGAAATGGGACAG 905
DB 889 GTGAGCGCTCATGAGGAGAAAGTTCAGGAGCAAGTAGTCTATAGAGAAATGGGCGAG 948
QY 906 TCACACGAGGAGACACCTCTGTTGGGCTCCAGCGGTTCTGATCTCAGTCCAGCCCAAGGCA 965
DB 949 TCACACGAGGAGACACCTCTGTTGGGCTCCAGCGGTTCTGATCTCAGTCCAGCCAGGCGC 1008
QY 966 GAGACAGCAGCTCTCTCAGCTCTCCGAAACAGTACCCCGAGTGGCCAGCGGAGCATGT 1025
DB 1009 GCGACAGCAGCTCTCTCAGCTCTCCGAAACAGTCTCCGAGTGGCCCGCGGAGCAGCATGT 1068

QY 1026 TTGACCATCCACCCCATCGAGCTCATCAAGGAAAGAGTAAGTCAGAGGAGTCCCTCT 1085
DB 1069 TTGAGCATCTCCCTCCCTGAGCTTGTGAAGTCAAGACTAAATCAAGAGAGTCTCTCT 1128
QY 1086 CTGACCTTTACAGGTTCCCTCTCTCCCTGAGCTTGTGTTGGCCCTCAGCTTTTGGATG 1145
DB 1129 CTGACCTGACCGGCTCCCTGCTCTCCCTGAGCTTGTGTTGGGCTTCACTTTTGGATG 1188
QY 1146 AGGTCTCAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAATTAAG 1205
DB 1189 AGTCTCAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAATTAAG 1248
QY 1206 AAAAGGTAC-----AAAAACAAGTAAACCAAGTGAAGAGAGAGG 1246
DB 1249 AAGAGGTACAAAATAAGCAGAAACAAATAAATAAATAAATAAATAAATAAATAAATAA 1308
QY 1247 CTTCGGGAGCTG-----TATTTGCGAGTTTGTGTTGGGTTTCTAAATPAA-----TATT 1296
DB 1309 CTGCTCTGACTGTTTATTTTATTTGAGCCATGATGATGATTTCTAAATAAAGCTCTCT 1368
QY 1297 CTTACAAAAGTATTTTATCTGTTATGCTGTTATGCTGTTTGCACAAACAAATTTAGAAAAAACA 1356
DB 1369 CTTACAAAAGTATTTTATCTGTTATGCTGTTTGCACAAACAAATTTAGAAAAAACA 1356
QY 1357 ACAAGCAAAACCTATCTTGGCAAAAAGGAA 1389
DB 1429 TTTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1461

RESULT 10
AF094521
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

AF094521
Homo sapiens MSE55-related protein (UBI) mRNA, complete cds.
AF094521.1 GI:3834632
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 765)
Alberts, A.S., Bouquin, N., Johnston, L.H. and Treisman, R.
Analysis of RhoA-binding proteins reveals an interaction domain
conserved in heterotrimeric G protein beta subunits and the yeast
response regulator protein Skn7
J. Biol. Chem. 273 (15), 8616-8622 (1998)
9826483
9535835
2 (bases 1 to 765)
Alberts, A.S.
MSE55-related Cdc42-binding protein
Unpublished
3 (bases 1 to 765)
Alberts, A.S.
Direct Submission
Submitted (24-SEP-1998) Cancer Center, University of California,
San Francisco, 2340 Sutter St., San Francisco, CA 94115, USA
Location/Qualifiers
1. .765
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="brain"
/dev_stage="fetal"
1. .765
/gene="UBI"
/note="Cdc42 binding protein; similar to Homo sapiens
serum protein MSE55; SwissProt Accession Number Q00587 and
unknown protein encoded by the sequence presented in
GenBank Accession Number AF001436; contains CRIB-motif"
/codon_start=1
/product="MSE55-related protein"



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 14:13:55 ; Search time 6693.77 Seconds
(without alignments)
11069.353 Million cell updates/sec

Title: US-09-806-276A-4
Perfect score: 2546
Sequence: 1 cggggctagccggagaccc.....cagctanaacgagggcagta 2546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2370.8	93.1	188359	9	AC006369	AC006369 Homo sapi
2	2165.6	85.1	2768	9	HS801810	AL136842 Homo sapi
3	1840	72.3	1892	9	BC019270	BC019270 Homo sapi
4	1215.8	47.8	93663	9	AC022409	AC022409 Homo sapi
5	1051.4	41.3	1118	9	AF164118	AF164118 Homo sapi
6	925.8	36.4	941	9	AF104857	AF104857 Homo sapi
7	904.6	35.5	216905	2	AC091332	AC091332 Mus muscu
8	895.6	35.2	292420	2	AC055705	AC055705 Mus muscu
9	769.8	30.2	1481	10	BC034714	BC034714 Mus muscu
10	761.8	29.9	765	9	AF094521	AF094521 Homo sapi
11	747.2	29.3	204019	2	AC090491	AC090491 Mus muscu
12	692.8	27.2	985	10	BC021409	BC021409 Mus muscu
13	393	15.4	432	11	G23498	G23498 human STS W
14	117.4	4.6	741	9	AF098290	AF098290 Homo sapi
15	117.4	4.6	1670	9	BC022337	BC022337 Homo sapi
16	117.4	4.6	1976	9	AF163840	AF163840 Homo sapi
17	117.4	4.6	1976	9	HSFA001436	AF001436 Human c10
18	117.4	4.6	2128	9	AK098292	AK098292 Homo sapi
19	117.4	4.6	78242	2	AP000483	AP000483 Homo sapi
20	117.4	4.6	171304	2	AC090297	AC090297 Homo sapi
21	117.4	4.6	218230	2	AC090390	AC090390 Homo sapi
22	117.4	4.6	233089	9	AC000353	AC000353 Homo sapi
23	115.8	4.5	1644	9	BC005406	BC005406 Homo sapi
24	115.2	4.5	149394	2	AP000828	AP000828 Homo sapi
25	108.4	4.3	1466	10	BC034884	BC034884 Mus muscu
26	108.4	4.3	203464	2	AC127271	AC127271 Mus muscu
27	107.6	4.2	51198	2	AP001279	AP001279 Homo sapi
28	88.8	3.5	200498	2	AP000944	AP000944 Homo sapi
29	64.4	2.5	110000	2	CEX11182_0	Z98857 Caenorhabdi
30	64.4	2.5	257703	3	CEX11182A	AL132904 Caenorhab
31	63.8	2.5	30726	2	AC117269	AC117269 Dictyoste
32	62.8	2.5	3364	14	TSP418778	AJ18778 Tomato sp
33	62.4	2.5	349980	6	AX344558	AX344558 Sequence
34	62	2.4	11812	6	AX251500	AX251500 Sequence
35	62	2.4	11812	6	AX278047	AX278047 Sequence
36	62	2.4	11812	6	AX323818	AX323818 Sequence
37	62	2.4	11812	6	AX347021	AX347021 Sequence
38	61.8	2.4	163547	2	AC111523	AC111523 Rattus no
39	61	2.4	131682	9	AL672277	AL672277 Human DNA
40	59.6	2.3	349980	6	AX344557	AX344557 Sequence
41	59.2	2.3	56772	3	AC006712	AC006712 Caenorhab
42	59	2.3	8079	6	AX356488	AX356488 Sequence
43	58.6	2.3	92971	9	AC023165	AC023165 Homo sapi
44	58.6	2.3	161523	2	AC084039	AC084039 Homo sapi
45	58.6	2.3	294914	3	AE003603	AE003603 Drosophila

ALIGNMENTS

RESULT 1
AC006369
LOCUS
Homo sapiens BAC clone RP11-278G12 from 2, complete sequence.
DEFINITION
AC006369
VERSION
AC006369.3 GI:5757529
KEYWORDS
HTG.
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 188359)
AUTHORS
Sulston,J.E. and Waterston,R.
TITLE
Toward a complete human genome sequence
JOURNAL
Genome Res. 8 (11), 1097-1108 (1998)

```

MEDLINE
PUBMED
99063792
9847074
2 (bases 1 to 188359)
Threiden,J., Elliot,G. and Modde,T.
TITLE
The sequence of Homo sapiens BAC clone RP11-278G12
JOURNAL
Unpublished
3 (bases 1 to 188359)
Waterston,R.H.
Direct Submission
Submitted (11-JAN-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 188359)
Waterston,R.H.
Direct Submission
Submitted (21-AUG-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 188359)
Waterston,R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 21, 1999 this sequence version replaced gi:4337247.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0278G12
-----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RP11-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(http://bacpac.med.buffalo.edu)
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
Actual start of this clone is at base position 1 of RP11-278G12;
actual end is at base position 188359 of RP11-278G12.
Location/Qualifiers
1. 188359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-278G12"
/clone_lib="RPC1-11"

FEATURES
Source
repeat_region 1171..1580
/rpt_family="L1"
repeat_region 1583..1878
/rpt_family="Retroviral"
repeat_region 1879..2102
/rpt_family="L1"
repeat_region 2174..2728
/rpt_family="L1"
repeat_region 2901..3020
/rpt_family="MERL_type"
repeat_region 3049..3201
/rpt_family="MIR"
repeat_region 3249..3375
/rpt_family="MIR"
repeat_region 3415..3465
/rpt_family="Mariner"
repeat_region 4160..4333
/rpt_family="MaLR"
repeat_region 4334..4640
/rpt_family="Alu"
repeat_region 4641..4858
/rpt_family="MaLR"
repeat_region 4859..5075
/rpt_family="MIR"
repeat_region 5076..5386
/rpt_family="Alu"
repeat_region 5387..5428
/rpt_family="MIR"
repeat_region 5903..6215
/rpt_family="L2"
repeat_region 6242..6301
/rpt_family="L2"
repeat_region 6302..7290
/rpt_family="Retroviral"
repeat_region 7291..7678
/rpt_family="L2"
repeat_region 7695..7758
/rpt_family="(CCCA)n"
repeat_region 7759..8057
/rpt_family="MaLR"
repeat_region 9868..9934
/rpt_family="L2"
repeat_region 10853..10821
/rpt_family="MIR"
repeat_region 10923..11219
/rpt_family="Alu"
repeat_region 11553..11708
/rpt_family="MIR"
repeat_region 12552..12968
/rpt_family="L1"
repeat_region 12985..13283
/rpt_family="Alu"
repeat_region 13284..13306
/rpt_family="(TAAA)n"
repeat_region 13307..13325
/rpt_family="Alu"
repeat_region 14650..14717
/rpt_family="(CAT)n"
repeat_region 14724..14965
/rpt_family="MIR"
repeat_region 15679..16057
/rpt_family="MaLR"
repeat_region 16334..16552
/rpt_family="Alu"
repeat_region 17921..18131
/rpt_family="MaLR"
repeat_region 18132..18330
/rpt_family="L1"
repeat_region 18910..19251
/rpt_family="MER73-group"
repeat_region 19252..19397
/rpt_family="L1"
repeat_region 19398..19620

```


Db 737 TGTCTCTGGTTGTGGTCAAGACTGGATGATGTAACGTGGCTCTCTAGAAAGCCCTCACTTGGC 796
Qy 241 CGTAACCTCAGGAAGTTCTCTTTGACCCCATCTCATTTTGAAGCCACTTCTGAAGCCAC 300
Db 797 CGTAACCTCAGGAAGTTCTCTTTGACCCCATCTCATTTTGAAGCCACTTCTGAAGCCAC 856
Qy 301 TTGAGAAAATGATGTGACAGTTCCTATCAAAAAGGATTTCAGAAACATATACCATCTGTG 360
Db 857 TTGAGAAAATGATGTGACAGTTCCTATCAAAAAGGATTTCAGAAACATATACCATCTGTG 916
Qy 361 AAGAAAAGTGGCCCTTTCTCCCGCTTGGCAAAATAGACATTCCTCAAAATCCAAATGCCAGC 420
Db 917 AAGAAAAGTGGCCCTTTCTCCCGCTTGGCAAAATAGACATTCCTCAAAATCCAAATGCCAGC 976
Qy 421 CAAGACCCCAATTTACCTGAAAGCAGCCCAATAAAGAAAGGAAAGAAATTTAAACTGTAG 480
Db 977 CAAGACCCCAATTTACCTGAAAGCAGCCCAATAAAGAAAGGAAAGAAATTTAAACTGTAG 1036
Qy 481 GGACATTTCTGCTCTCGATATGATGATGCTCCCGCTTGGAGACTTTCCGCACACCATCCA 540
Db 1037 GGACATTTCTGCTCTCGATATGATGATGCTCCCGCTTGGAGACTTTCCGCACACCATCCA 1096
Qy 541 CATTGGCAAGAGGCCAGCAGATGCTTTTGGAGATATTCCTTTCTTCAAGGGAACTA 600
Db 1097 CATTGGCAAGAGGCCAGCAGATGCTTTTGGAGATATTCCTTTCTTCAAGGGAACTA 1156
Qy 601 CGAGCTTTTACCTGGAACACAGGAGAAAGCACACCTGGGCCAGTTCCCTGGGCATATCA 660
Db 1157 CGAGCTTTTACCTGGAACACAGGAGAAAGCACACCTGGGCCAGTTCCCTGGGCATATCA 1216
Qy 661 GTTCTTCCGGGCCAACAGCACCCTCGGACTCTGTGTCAGAAACGCCCTCCCGGTGCT 720
Db 1217 GTTCTTCCGGGCCAACAGCACCCTCGGACTCTGTGTCAGAAACGCCCTCCCGGTGCT 1276
Qy 721 CAAAATGCCATCTCCCTCCGACCATTTGAGAGATCCCAAGCTCTCATGTTGCCCTTAT 1336
Db 1277 CAAAATGCCATCTCCCTCCGACCATTTGAGAGATCCCAAGCTCTCATGTTGCCCTTAT 1396
Qy 781 GTCACCACTGACATTTAAATTTCCAAACAGGAGTCCCTCGGCCAGCAAGCTGCCAGGCT 840
Db 1337 GTCACCACTGACATTTAAATTTCCAAACAGGAGTCCCTCGGCCAGCAAGCTGCCAGGCT 1396
Qy 841 TAGCTCGAGCCCGTCTATGGAGGAAAAGCTCAGGAGAAAAGCTGTGTTGGAGATGG 900
Db 1397 TAGCTCGAGCCCGTCTATGGAGGAAAAGCTCAGGAGAAAAGCTGTGTTGGAGATGG 1456
Qy 901 GACAGTCCACAGGAGACACCTCTGTGGGCTCCAGCGGTTCTGCATCTCAGTCCAGCCA 960
Db 1457 GACAGTCCACAGGAGACACCTCTGTGGGCTCCAGCGGTTCTGCATCTCAGTCCAGCCA 1516
Qy 961 AGGCAGAGACGCACTCTCCAGCCTCTCCGAAACAGTACCCCGACTGGCCAGCCGAGGA 1020
Db 1517 AGGCAGAGACGCACTCTCCAGCCTCTCCGAAACAGTACCCCGACTGGCCAGCCGAGGA 1576
Qy 1021 CATGTTTGACCATCCACCCCATCGGAGCTCATCAAGGGAAGACTAAGTCAGAGGAGTC 1080
Db 1577 CATGTTTGACCATCCACCCCATCGGAGCTCATCAAGGGAAGACTAAGTCAGAGGAGTC 1636
Qy 1081 CCTCTGACCTTACAGGTTCCCTCTCTCCCTGACGTTGATCTGGGCCCTTCACTTTT 1140
Db 1637 CCTCTGACCTTACAGGTTCCCTCTCTCCCTGACGTTGATCTGGGCCCTTCACTTTT 1696
Qy 1141 GGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1697 GGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1756
Qy 1201 GGGTAAAGGTACAAAAACAACTAACACAGTGAAGAGAGGCTTCCGGAGCTGTA 1260
Db 1757 GGGTAAAGGTACAAAAACAACTAACACAGTGAAGAGAGGCTTCCGGAGCTGTA 1816
Qy 1261 TTTGACGTTTGTGTTGTTTCTTAAATAATATCTTTTACAAAGTATTTTTTACCTGT 1320
Db 1817 TTTGACGTTTGTGTTGTTTCTTAAATAATATCTTTTACAAAGTATTTTTTACCTGT 1876

Qy 1321 TATGCCCTGTTTGCAGAAAACAATTTAGAAAAAACAACAAAGCAAAACCTATCTTGCAA 1380
Db 1877 TATGCCCTGTTTGCAGAAAACAATTTAGAAAAAACAACAAAGCAAAACCTATCTTGCAA 1936
Qy 1381 AAAAAGAAAGTGTGAGTCAGAGCCCATTTTCAGGAGGCATTTGTGTGTTGCGCTCACATAT 1440
Db 1937 AAAAAGAAAGTGTGAGTCAGAGCCCATTTTCAGGAGGCATTTGTGTGTTGCGCTCACATAT 1996
Qy 1441 TGTGTCAGACACACAGAAATCTGGCTTGGCCAGGATTGGCACTAGCTATGAAGGCGTG 1500
Db 1997 TGTGTCAGACACACAGAAATCTGGCTTGGCCAGGATTGGCACTAGCTATGAAGGCGTG 2056
Qy 1501 AGCGAGTCACATTAAGGAACTTCACGGAACTTTATAGCACTCCGACATTTTCTGAGCAAG 1560
Db 2057 AGCGAGTCACATTAAGGAACTTCACGGAACTTTATAGCACTCCGACATTTTCTGAGCAAG 2116
Qy 1561 AGGAAGTCACAAATTTATTTAACACCTAAGCCCTTTTCTAGACTCTTTTCTATATATGCT 1620
Db 2117 AGGAAGTCACAAATTTATTTAACACCTAAGCCCTTTTCTAGACTCTTTTCTATATATGCT 2176
Qy 1621 TAGGCTCACCATAGCGAATTCCTCAGTGTAAACCTTTTCTGTTTTCACATTTGAACCTT 1680
Db 2177 TAGGCTCACCATAGCGAATTCCTCAGTGTAAACCTTTTCTGTTTTCACATTTGAACCTT 2236
Qy 1681 ATGGGTTTGGGGATTTTCTGTAGTCTTATATATATATATATATATATATATATATGCT 1740
Db 2237 ATGGGTTTGGGGATTTTCTGTAGTCTTATATATATATATATATATATATATATATGCT 2296
Qy 1741 AAAATTTTGACTGTCAGCTACATGTTGGTAAGACACAGGCAAAAGTATTTACTGTAAGTAA 1800
Db 2297 AAAATTTTGACTGTCAGCTACATGTTGGTAAGACACAGGCAAAAGTATTTACTGTAAGTAA 2356
Qy 1801 TTATTTTAAAGTTAAATATATTTTACGTGCTTGGCTTTTATTCGAGAGTCTACA 1860
Db 2357 TTATTTTAAAGTTAAATATATTTTACGTGCTTGGCTTTTATTCGAGAGTCTACA 2416
Qy 1861 TTTATATAGATTTTACATCAGATGTTGTCACCTTATTTCCATTTGGGATTTCCATTTAGCTG 1920
Db 2417 TTTATATAGATTTTACATCAGATGTTGTCACCTTATTTCCATTTGGGATTTCCATTTAGCTG 2476
Qy 1921 TGTATGCTGCTGTTTGGAAAAAGTGTATTCATCACTTAGTTTTCTTCTCATCTGTTATC 1980
Db 2477 TGTATGCTGCTGTTTGGAAAAAGTGTATTCATCACTTAGTTTTCTTCTCATCTGTTATC 2536
Qy 1981 ATACTTTTAAACAGCAACCAATTAACGATTTGTAAGTGTAAAGGACAGGTTTACTCATGAT 2040
Db 2537 ATACTTTTAAACAGCAACCAATTAACGATTTGTAAGTGTAAAGGACAGGTTTACTCATGAT 2596
Qy 2041 GCTTCTGACAGACTGTGGGCTTACACCATATGTTATTTGGAATATAGTATTTAGT 2100
Db 2597 GCTTCTGACAGACTGTGGGCTTACACCATATGTTATTTGGAATATAGTATTTAGT 2656
Qy 2101 ACAGTACATCTTGCATTACATAGTACTTCAAGCAACACAAATAAAAAGTAAATGATAA 2160
Db 2657 ACAGTACATCTTGCATTACATAGTACTTCAAGCAACACAAATAAAAAGTAAATGATAA 2716
Qy 2161 GTGAACCTGCTGTTTATAGTAAATTAACAGACCATTAAGAGATAAGTATAG 2212
Db 2717 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 2768

RESULT 3
BC019270
LOCUS
DEFINITION
BC019270
ACCESSION
VERSION
BC019270.1
KEYWORDS
MGC.
SOURCE
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

BC019270 1892 bp mRNA linear PRI 19-DEC-2001
Homo sapiens, clone MGC:3481 IMAGE:2823261, mRNA, complete cds.
BC019270.1 GI:17939542

REFERENCE	1 (bases 1 to 1892)	
AUTHORS	Strausberg R.	
TITLE	Direct Submission	
JOURNAL	Submitted (13-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC project URL: http://mgc.nci.nih.gov	
COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Ruben Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalrus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.	
FEATURES	Location/Qualifiers	
source	1..1892	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="MGC:3481 IMAGE:2823261"	
	/tissue_type="Lung, small cell carcinoma"	
	/clone_lib="NIH_MGC_7"	
	/lab_host="DH10B-R"	
	/note="Vector: pOTB7"	
	438..1202	
	/codon_start=1	
	/product="Unknown (protein for MGC:3481)"	
	/protein_id="AAH19270.1"	
	/db_xref="GI:17939543"	
	/translation="MPAKTPYVLAANNKKKPKLRLSPDMISPLDGRHTTHI GREGHDVDFSLFQGNVLELPQNOEKHLGFPNHEFRANSTSDSVFTETPSV LKNALSLPTIGSQALMLPLSPVTFNSKQSFPAKLPLRSLCEPVMEEKAQKSL ENGTWQDTSNGSSASQSSQGRDSSHSSLSQYPDWPAEDMFDHTPCELKIGKT KSESLSDLTGSLSLSLQLDLGFLDLVNMKNK"	
BASE COUNT	525 a 462 c 434 g 471 t	
ORIGIN		
Query Match	72.3%; Score 1840; DB 9; Length 1892;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1840:	Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 CGGGGCTAGCCCGGAGACCGGGCCACCGGCTGGGGCGCTTCACGCCGCTCTCGAGCGG 60	
Db		
Qy	26 CGGGGCTAGCCCGGAGACCGGGCCACCGGCTGGGGCGCTTCACGCCGCTCTCGAGCGG 85	
Db		
Qy	61 ATAATCGGTGAGCAGGACACCGCGGAGACTCGGCTGGATCTCGGCACAGCGGCGAG 120	
Db		
Qy	86 ATAATCGGTGAGCAGGACACCGCGGAGACTCGGCTGGATCTCGGCACAGCGGCGAG 145	
Db		
Qy	121 GATTGCGTGGCCCGGGAGGCGCGGCGGCTGGATCTCGGCACAGCGGCGGCTT 180	
Db		
Qy	146 GATTGCGTGGCCCGGGAGGCGCGGCGGCTGGATCTCGGCACAGCGGCGGCTT 205	
Db		
Qy	181 TCTCTGCTGTGTGTCAGACTGTGATGATGATGATGATGATGATGATGATGATGATGAT 240	
Db		
Qy	206 TCTCTGCTGTGTGTCAGACTGTGATGATGATGATGATGATGATGATGATGATGATGAT 265	
Db		
Qy	241 CGTAACCTCAGGAAGGTTCTCTTTGACCCCATCTCTTCTGAGCCACTTCTGAAGCCAC 300	
Db		
Qy	266 CGTAACCTCAGGAAGGTTCTCTTTGACCCCATCTCTTCTGAGCCACTTCTGAAGCCAC 325	
Db		

Qy	301 TTGAGAAAATGATGACAGTTCCTATCAAAAAGGATTTCAGAAACATATACCATCTGTG 360	
Db		
Qy	326 TTGAGAAAATGATGACAGTTCCTATCAAAAAGGATTTCAGAAACATATACCATCTGTG 385	
Db		
Qy	361 AAGAAGTGGCCCTTTCTCCCGCTTGCAGAAATAGACATCTCAATTTCCAAATGCCAGC 420	
Db		
Qy	386 AAGAAGTGGCCCTTTCTCCCGCTTGCAGAAATAGACATCTCAATTTCCAAATGCCAGC 445	
Db		
Qy	421 CAAGACCCCAATTTACCTGAAGAGCAGCAATAAAGAAAGGAAAGAAATTTAAACATGAG 480	
Db		
Qy	446 CAAGACCCCAATTTACCTGAAGAGCAGCAATAAAGAAAGGAAAGAAATTTAAACATGAG 505	
Db		
Qy	481 GGACATCTCTCTCTGATATGATCAGTCCCGCTTGGAGACTTTCCGACACCATCCA 540	
Db		
Qy	506 GGACATCTCTCTCTGATATGATCAGTCCCGCTTGGAGACTTTCCGACACCATCCA 565	
Db		
Qy	541 CATGGCAAGAGGGCCAGCAGATCTCTTGGAGATATTTCTTCTTCAAGGGAACATA 600	
Db		
Qy	566 CATGGCAAGAGGGCCAGCAGATCTCTTGGAGATATTTCTTCTTCAAGGGAACATA 625	
Db		
Qy	601 CGAGCTTTTACCTGGAAACCCAGGAGAAACACACCTGGGCCAGTTCCTGGGCATAATGA 660	
Db		
Qy	626 CGAGCTTTTACCTGGAAACCCAGGAGAAACACACCTGGGCCAGTTCCTGGGCATAATGA 685	
Db		
Qy	661 GTTCTTCCGGCCACAGCAGCTCGGACTCTGTGTTTTCAGAAACGCCCTCCCGGTGCT 720	
Db		
Qy	686 GTTCTTCCGGCCACAGCAGCTCGGACTCTGTGTTTTCAGAAACGCCCTCCCGGTGCT 745	
Db		
Qy	721 CAAAAATGCCATCTCCCTCCCGACCAATTTGGAGATCCCAAGCTCTCATGTGCGCTTATT 780	
Db		
Qy	746 CAAAAATGCCATCTCCCTCCCGACCAATTTGGAGATCCCAAGCTCTCATGTGCGCTTATT 805	
Db		
Qy	781 GTACACAGTGACATTTAATTCACAAAGAGGAGTCTCTCGGGCCAGCAAGAGTCCCGAGGCT 840	
Db		
Qy	806 GTACACAGTGACATTTAATTCACAAAGAGGAGTCTCTCGGGCCAGCAAGAGTCCCGAGGCT 865	
Db		
Qy	841 TAGCTGCGAGCCGCTCATGGAGGAAAAGCTCAGGAGAAAAGAGTCTGTGGAGATGG 900	
Db		
Qy	866 TAGCTGCGAGCCGCTCATGGAGGAAAAGCTCAGGAGAAAAGAGTCTGTGGAGATGG 925	
Db		
Qy	901 GACAGTCCACAGGAGGAGACACTCGTGGGGCTCCAGCGGTTCTGCATCTCAGTCCAGGCCA 960	
Db		
Qy	926 GACAGTCCACAGGAGGAGACACTCGTGGGGCTCCAGCGGTTCTGCATCTCAGTCCAGGCCA 985	
Db		
Qy	961 AGCGAGAGACGCCACTCTCTCAGCGCTTCCGGAACAGTACCCGACTGGCCAGCCGAGGA 1020	
Db		
Qy	986 AGCGAGAGACGCCACTCTCTCAGCGCTTCCGGAACAGTACCCGACTGGCCAGCCGAGGA 1045	
Db		
Qy	1021 CATGTTTGACCATCCCAATCGGAGCTCATCAAGGAAAGAGTCAAGTCAGAGGAGTC 1080	
Db		
Qy	1046 CATGTTTGACCATCCCAATCGGAGCTCATCAAGGAAAGAGTCAAGTCAGAGGAGTC 1105	
Db		
Qy	1081 CCTCTGACCTTACAGGTTCCCTCTCTCCCTGAGCTTGATCTTGGGCGCTCAGCTTTT 1140	
Db		
Qy	1106 CCTCTGACCTTACAGGTTCCCTCTCTCCCTGAGCTTGATCTTGGGCGCTCAGCTTTT 1165	
Db		
Qy	1141 GGATGAGTGTCTGAATGATTAATGATTAATTAAGTAACAAGATGCCAACTTTTTCCTTT 1200	
Db		
Qy	1166 GGATGAGTGTCTGAATGATTAATGATTAATTAAGTAACAAGATGCCAACTTTTTCCTTT 1225	
Db		
Qy	1201 GGGTAAAGGTACAAAACAAACTAACACAGTTTGAAGAGAAAGGCTTCGGAGCTGTA 1260	
Db		
Qy	1226 GGGTAAAGGTACAAAACAAACTAACACAGTTTGAAGAGAAAGGCTTCGGAGCTGTA 1285	
Db		
Qy	1261 TTTGAGTGTGTGGGTTTTCTTAAATAATATTTCTTACAAAGTATTTTTTACCTTCT 1320	
Db		
Qy	1286 TTTGAGTGTGTGGGTTTTCTTAAATAATATTTCTTACAAAGTATTTTTTACCTTCT 1345	
Db		
Qy	1321 TATGCCCTGTGTCAAAACAAATTTAGAAAAAACCAACAAAGCAAAACCTATCTGGGCAA 1380	
Db		
Qy	1346 TATGCCCTGTGTCAAAACAAATTTAGAAAAAACCAACAAAGCAAAACCTATCTGGGCAA 1405	
Db		

```

QY 1381 AAAAGGAAGTGTAGTCAGAGCCCATTTTCAGGAGGCAATTTGGTGTATGTCGGCTCACATAT 1440
Db 1406 AAAAGGAAGTGTAGTCAGAGCCCATTTTCAGGAGGCAATTTGGTGTATGTCGGCTCACATAT 1465
QY 1441 TGTTCGAGACACACAAGAAATCTGGCTTGGCCAGGATTTGGCACTAGCTATGAAGGGCTG 1500
Db 1466 TGTTCGAGACACACAAGAAATCTGGCTTGGCCAGGATTTGGCACTAGCTATGAAGGGCTG 1525
QY 1501 AGCGAGTCACATTAAGGAATCTCAGGAACTTTATAGCACTCCGACATTTCTGAGCAAG 1560
Db 1526 ACCAGTCACATTAAGGAATCTCAGGAACTTTATAGCACTCCGACATTTCTGAGCAAG 1585
QY 1561 AGGAAGTCAAAATTTATTTAACACCTTAAGCCTTTTGTAGACTCTTTTCTATATATGCT 1620
Db 1586 AGGAAGTCAAAATTTATTTAACACCTTAAGCCTTTTGTAGACTCTTTTCTATATATGCT 1645
QY 1621 TAGCTCACCATACGGAATCTCCAGGTTTAAACATTTTCTGTTTTCACATTTGAACCTT 1680
Db 1646 TAGCTCACCATACGGAATCTCCAGGTTTAAACATTTTCTGTTTTCACATTTGAACCTT 1705
QY 1681 ATGGGTTTTGGGATTTTCTGTAGTTCTTATATATATATATATATATATATATATATGCT 1740
Db 1706 ATGGGTTTTGGGATTTTCTGTAGTTCTTATATATATATATATATATATATATATATGCT 1765
QY 1741 AAAATTTTGTCTCAGTACATGTTGGTAAAGACACAGGCAAGATTAATGTAACCTAAG 1800
Db 1766 AAAATTTTGTCTCAGTACATGTTGGTAAAGACACAGGCAAGATTAATGTAACCTAAG 1825
QY 1801 TTATTTTAAAGTTAAATATATTTTACGTGCTCTTGGC 1840
Db 1826 TTATTTTAAAGTTAAATATATTTTACGTGCTCTTGGC 1865

RESULT 4
AC022409 93663 bp DNA linear PRI 23-AUG-2001
LOCUS Homo sapiens chromosome 19 clone CTB-175P5, complete sequence.
AC022409
VERSION AC022409.7 GI:15281220
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 93663)
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE
AUTHORS 2 (bases 1 to 93663)
TITLE DOE Joint Genome Institute.
JOURNAL Direct Submission
REFERENCE
AUTHORS 3 (bases 1 to 93663)
TITLE Submitted (03-FEB-2000) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE
AUTHORS 4 (bases 1 to 93663)
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Submitted (08-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Aug 23, 2001 this sequence version replaced gi:11120761.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >40 99.4% of Sequence;
Estimated Total Number of Errors is 0.2.
STS Content:
SHGC-150482 G60581.
Location/Qualifiers

```

```

source 1. 93663
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-175P5"
BASE COUNT 27383 a 19719 c 19712 g 26849 t
ORIGIN

Query Match 47.8%; Score 1215.8; DB 9; Length 93663;
Best Local Similarity 89.4%; Pred. No. 2.2e-285;
Matches 1407; Conservative 0; Mismatches 137; Indels 29; Gaps 8;

QY 283 AGCCACATCTTGAAGCCACTTTGAGAAAATGATGACAGTTCCTATCAAAAAGGATTCAG 342
Db 13855 AGCCACATCTTGAAGCCACTTTGAGAAAATTTGACAGTTCCTACGAAAAGATTTCTTCAG 13914
QY 343 AAACATATACCATCTGTGAAGAAGTGGCGCTTTCTCCGCTTTGAAAATATACACATTTCTC 402
Db 13915 AAACATATACCATCTGTGAAGAAGTGGCGCTTTCTCCGCTTTGAAAATATACACATTTCTC 13974
QY 403 AAATTTCCAAATGCCAGCCAGCCCAATTTACCTGAAAGCAGGCCAATAACAAGAAAGG 462
Db 13975 AAATTTCCAAATGCCAGCCAGCCCAATTTACCTGAAAGCAGGCCAATAACAAGAAAGG 14034
QY 463 AAAGAAATTTAAAGCTGAGGACATTTCTGTCTCTGTATGATGATCAGTCCCGCGCTTGGAGA 522
Db 14035 AAAGAAATTTAAAGCTGAGGACATTTCTGTCTCTGTATGATGATCAGTCCCGCGCTTGGAGA 14092
QY 523 CTTTGGCCACACCATCCACATTTGCCAAAGAGGCC-AGCAGATGCTTTTGGAGATATTT 581
Db 14093 CTTTGGCCACACCATCCACATTTGCCAAAGAGGCCAAGCATGACGTCTTTTGGAGATATTT 14152
QY 582 CCTTTCTCAAGGAACTACGAGCTTTTACCTGAAAACACAGGAGAACACACCTGGGCC 641
Db 14153 CCTTTCTCAAGGAACTACGAGCTTTTACCTGAAAACACAGGAGAACACACCTGGGCC 14212
QY 642 AGTTCCTTGGCCATAATGAGTTCTTCCGGGCCAACAGCAGCTCGGAGCTGTGTTCACAG 701
Db 14213 AGTTCCTTGGCCATAATGAGTTCTTCCGGGCCAACAGCAGCTCGGAGCTGTGTTCACAG 14272
QY 702 AAACGCCCTCCCGGCTGCTCAAAATGCATCTCCCTCCGACCATTTGGAGGATCCCAAG 761
Db 14273 AAACGCCCTCCCGGCTGCTCAAAATGCATCTCCCTCCGACCATTTGGAGGATCCCAAG 14332
QY 762 CTCTCATGTTGCCCTTATTGTACAGTGTACATTTAATTCAAAACAGAGTTCCTTCGGCC 821
Db 14333 CTCTCATGTTGCCCTTATTGTACAGTGTACATTTAATTCAAAACAGAGTTCCTTCGGCC 14392
QY 822 CAGCAAAAGTGCCTCCAGGCTTAGCTCGGAGCCCTGATGGAGGAAAAGCTCAGGAGAAA 881
Db 14393 CAGCAAAAGTGCCTCCAGGCTTAGCTCGGAGCCCTGATGGAGGAAAAGCTCAGGAGAAA 14452
QY 882 CGAGTCTGTGGAGAAATGGGACAGTCCACAGGAGACACCTCTCGTGGGGTTCACAGCGTT 941
Db 14453 CGAGTCTGTGGAGAAATGGGACAGTCCACAGGAGACACCTCTCGTGGGGTTCACAGCGTT 14512
QY 942 CTGCATCTCAGTCCAGCAAGGACAGACAGCCACTCTCTCCAGCCTGTCCGAACAGTACC 1001
Db 14513 CTGCATCTCAGTCCAGCAAGGACAGACAGCCACTCTCTCCAGCCTGTCCGAACAGTACC 14572
QY 1002 CGCAATGGCCAGCCAGGACATGTTGACCATCCCAACCCCATCGAGCTCATCAAGGAA 1061
Db 14573 CGCAATGGCCAGCCAGGACATGTTGACCATCCCAACCCCATCGAGCTCATCAAGGAA 14632
QY 1062 AGACTAAGTCAGAGGAGTCCCTCTCTGACCTTACAGGTTCCCTCTCTCCCTGAGCTTG 1121
Db 14633 AGACTAAGTCAGAGGAGTCCCTCTCTGATCTTACAGGTTCCCTCTCTCCCTGAGCTTG 14692
QY 1122 ATCTTGGGCCCTCAGTCTTTTGGATGAGGTCGTAATGTAATGATAAAATTAAGTAACAG 1181
Db 14693 ATCTTGGGCCCTCAGTCTTTTGGATGAGGTCGTAATGTAATGATAAAATTAAGTAACAG 14752
QY 1182 ATGCCAAACCTTTTCTCTTTGGGGTAAAGGTACAAAAC----AACTAACACAGTTG 1236

```

University of Virginia, West Complex Building, Charlottesville, VA 22908, USA

Db	481	GAGAAAAGCAGTCTGTTGGAGATGGACAGTCCACAGGAGACACCTCGTGGGGTCC	540		
Qy	935	AGCGGTTCTGCTCAGTCCAGCCAGGAGAGAGCCACTCTCCAGGCTGTCCGAA	994		
Db	541	AGCGGTTCTGCTCAGTCCAGCCAGGAGAGAGCCACTCTCCAGGCTGTCCGAA	600		
Qy	995	CAGTCCCGGACTGGCCAGCGAGGACATGTTTGACCATCCACCCCATGGAGCTATC	1054		
Db	601	CAGTCCCGGACTGGCCAGCGAGGACATGTTTGACCATCCACCCCATGGAGCTATC	660		
Qy	1055	AAGGGAAGACTAAGTCAGAGGAGTCCCTCTCTGACCTTACAGGTTCCCTCTCCCTG	1114		
Db	661	AAGGGAAGACTAAGTCAGAGGAGTCCCTCTCTGACCTTACAGGTTCCCTCTCCCTG	720		
Qy	1115	CAGCTTGATCTTGGGCGCTCACTTTTGGATGAGTGTGATGTAATGGATAAAATAAG	1174		
Db	721	CAGCTTGATCTTGGGCGCTCACTTTTGGATGAGTGTGATGTAATGGATAAAATAAG	780		
Qy	1175	TACAAAGATGCCAATTTTTTCCCTTTGGGTAAGGTACAAAAACAACTAACCCAGT	1234		
Db	781	TACAAAGATGCCAATTTTTTCCCTTTGGGTAAGGTACAAAAACAACTAACCCAGT	840		
Qy	1235	TGAAGAGAGGCTTCCGAGCTGATTTGACGTTTGTGTTGGTTTCTAAATAATA	1294		
Db	841	TGAAGAGAGGCTTCCGAGCTGATTTGACGTTTGTGTTGGTTTCTAAATAATA	900		
Qy	1295	TTCTTACAAAGTATTTTTTACCTGTTATGCGCTTTTGCAGTTTGTGTTGGTTTCTAAATAATA	1354		
Db	901	TTCTTACAAAGTATTTTTTACCTGTTATGCGCTTTTGCAGTTTGTGTTGGTTTCTAAATAATA	959		
Qy	1355	CAACAAAGCAAACTATCTTGGCAAAAAAGGAAGTGTAGTACAG-AGCCCATTTTCAGGA	1413		
Db	960	RACCAACAAACCTATC-TGGCAAAAAAGGAAGTGTAGTACAGAGCCCATTTTCAGGA	1018		
Qy	1414	GCAATGTTGATGTTTGGCTCACATATTTTGCAGACACACAGAAATCTGCTTGGCC	1473		
Db	1019	GCAATGTTGATGTTTGGCTCACATATTTTGCAGACACACAGAAATCTGCTTGGCC	1077		
Qy	1474	AGGATGGCAC-TAGCTATGAGGGCTGAGCGAGTACACTT	1513		
Db	1078	AGGATGGCAC-TAGCTATGAGGGCTGAGCGAGTACACTT	1118		
RESULT 6					
LOCUS	AF104857				
DEFINITION	Homo sapiens Cdc42 effector protein 3 mRNA	linear	PRI 09-JAN-2001		
ACCESSION	AF104857				
VERSION	AF104857.1	GI:4324453			
KEYWORDS					
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE					
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
JOURNAL	Hirsch, D.S., Pirone, D.M. and Burbelo, P.D.				
MEDLINE	A new family of Cdc42 effector proteins, CEPs, function in				
PUBMED	fibroblast and epithelial cell shape changes				
REFERENCE	J. Biol. Chem. 276 (2), 875-883 (2001)				
AUTHORS	2 (bases 1 to 941)				
TITLE	Thomas, D.S. and Burbelo, P.D.				
JOURNAL	Unique domains define a new family of Cdc42 effector proteins				
AUTHORS	(CEPs)				
TITLE	Unpublished				
JOURNAL	3 (bases 1 to 941)				
AUTHORS	Thomas, D.S. and Burbelo, P.D.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-NOV-1998) Lombardi Cancer Center, Georgetown				
	University Medical Center, EG16 New Research Bldg, 3970 Reservoir				
	Road, NW, Washington, DC 20007-2197, USA				

FEATURES	Location/Qualifiers				
source	1..941				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE Consortium ID 488206"				
	/sex="female"				
	/tissue_type="pregnant uterus"				
CDS	61..825				
	/note="CEP3"				
	/codon_start=1				
	/product="Cdc42 effector protein 3"				
	/protein_id="AA16888.1"				
	/db_xref="GI:4324454"				
	/translation="MPAKTPIYLKAANNKKGKELRDLISPDMSPLPGLDFRHTTHI				
	LKNGDHDVFGNSFLQNGYELLPGNQKHAHLQFPFHNFRANSTSDSVTFPTSPV				
	LKNAISLPTIGGSQALMLPLLSQVFNSKQESFGPAKLRLSCPEVMEERAKQKSSLL				
	ENGTVHQDTSWSSGSASQSGRDSLSLSSEYQPDWPAEDMFDPHTPCELKIGKT				
	KSESLDLTGLSLSLQDLGLPSLLDEVLYNMDKNK"				
BASE COUNT	259 a 256 c 213 g 213 t				
ORIGIN					
Query Match	36.4%; Score 925.8; DB 9; Length 941;				
Best Local Similarity	99.7%; Pred. No. 5.9e-215;				
Matches	938; Conservative 0; Mismatches 2; Indels 1; Gaps 1;				
Qy	353	CATCTGTGAGAAAGTGGCCCTTTCTCCGCTTGCAAAATAGACATTTCAAAATTCAAA	412		
Db	1	CATCTGTGAGAAAGTGGCCCTTTCTCCGCTTGCAAAATAGACATTTCAAAATTCAAA	60		
Qy	413	ATGCCAGCAAGACCCCAATTTACCTGAAGAGCCCAATTAACAGAAAGGAAAGAAATTT	472		
Db	61	ATGCCAGCAAGACCCCAATTTACCTGAAGAGCCCAATTAACAGAAAGGAAAGAAATTT	472		
Qy	473	AACTGAGGGACATTTCTCTCTGATATGATCAGTCCCGCTTGAGAGACTTTTCCGCCAC	532		
Db	121	AACTGAGGGACATTTCTCTCTGATATGATCAGTCCCGCTTGAGAGACTTTTCCGCCAC	180		
Qy	533	ACCATCCACATTTGCCAAGAGGCGGACGACATGCTTTGGAGATATTTCTTTTCTCAA	592		
Db	181	ACCATCCACATTTGCCAAGAGGCGGACGACATGCTTTGGAGATATTTCTTTCTCAA	240		
Qy	593	GGGAACACGAGCTTTTACCTGGAAACACAGGAGAAAGACACCTGGCCAGTTCCTTGGG	652		
Db	241	GGGAACACGAGCTTTTACCTGGAAACACAGGAGAAAGACACCTGGCCAGTTCCTTGGG	300		
Qy	653	CATAATCAGTCTTCGGGCCAACAGACCTCGGACTCTGTGTTTCAGAAACGCCCTCC	712		
Db	301	CATAATCAGTCTTCGGGCCAACAGACCTCGGACTCTGTGTTTCAGAAACGCCCTCC	360		
Qy	713	CCGCTGCTCAAAAATGCATCTCCCTCCGACCATTTGGAGATCCCAAGCTCTCATGTTG	772		
Db	361	CCGCTGCTCAAAAATGCATCTCCCTCCGACCATTTGGAGATCCCAAGCTCTCATGTTG	420		
Qy	773	CCCTTATGTCACCTGACATTTAATCCAAACAGGAGTCTTCCGGCCAGCAAGCTG	832		
Db	421	CCCTTATGTCACCTGACATTTAATCCAAACAGGAGTCTTCCGGCCAGCAAGCTG	480		
Qy	833	CCGAGGCTTAGCTCGGAGCCGCTCATGAGGAGAAAGCTCAGGAGAAAGCAGTCTGTTG	892		
Db	481	CCGAGGCTTAGCTCGGAGCCGCTCATGAGGAGAAAGCTCAGGAGAAAGCAGTCTGTTG	540		
Qy	893	GAGATGGGACAGTCCACCCAGGAGACACCTCGTGGGCTCCAGCGGTTCTGCATCTCAG	952		
Db	541	GAGATGGGACAGTCCACCCAGGAGACACCTCGTGGGCTCCAGCGGTTCTGCATCTCAG	600		
Qy	953	TCCAGCAAGGAGAGACAGCCACTTCTCCAGCTGTCCGAACAGTACCCGAGCTGGCCA	1012		
Db	601	TCCAGCAAGGAGAGACAGCCACTTCTCCAGCTGTCCGAACAGTACCCGAGCTGGCCA	660		
Qy	1013	GCCGAGGACATGTTTGACCATCCCAACCCATCCGAGCTCATCAAGGAGAAAGCTAAGTCA	1072		
Db	661	GCCGAGGACATGTTTGACCATCCCAACCCATCCGAGCTCATCAAGGAGAAAGCTAAGTCA	720		

Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, D., Trefse, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 20, 2002 this sequence version replaced gi:14209739.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13294

* consists of 9 contigs, and the order in this sequence record is
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

191022 * 210905Z Conting Of 22004 Sp in 2093
FEATURES
source
Location/Qualifiers
1. .216905
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="Rp23-12k1"
/clone_lib="RPCI-23 Female Mouse BAC"
misc_feature
1. .78339
/note="assembly_fragment
clone_end:SP6
vector_side:left"
78440..79696
/note="assembly_fragment"
misc_feature
79797..83334
/note="assembly_fragment"
misc_feature
83435..89559
/note="assembly_fragment"
misc_feature
89560..97693

```



```
misc_feature /note="assembly_fragment"
97794..113828 /note="assembly_fragment"
misc_feature 113929..1139314 /note="assembly_fragment"
misc_feature 139415..190921 /note="assembly_fragment"
misc_feature 191022..216905 /note="assembly_fragment"
clone_end:17
vector_side:right"
BASE COUNT 62522 a 44706 c 45987 g 62878 t 812 others
ORIGIN
Query Match 35.5%; Score 904.6; DB 2; Length 216905;
Best Local Similarity 69.2%; Pred. No. 2.4e-209;
Matches 1503; Conservative 0; Mismatches 554; Indels 116; Gaps 15;
QY 175 CCGGTTTTCCTGGTGGTCAAGACTGGATGATGAAGTAACTGGCTCTCTAGGAAGCCCTCA 234
DB 205547 CAGGTTTCTCTGAGCTGTGGAAGGACTCGATAATGTAAGTAACTGGCTCTCGGAAGACCCG 205488
QY 235 CTTGGCCGTAACCTCAGGAAGTCTCTTTGACCCCATCTCATTTGGAAGCCACTTCTGA 294
DB 205487 CGTGGCCCTCGCTGAGCAAGATCTCCAGACCCCTGCTCTGCTTTTATAGACCACTTGAGA 205428
QY 295 AGCCACTTCAGAAAATGATGACAGTTCTCTATCAAAAAGGA---TTCAGAAAACATATA 351
DB 205427 AG-----AATGATGTGACAAATCTTATGAAAAGGACGCTTTTGGAGGTGATAG 205379
QY 352 CCATCTGTGAAGAAAGTGGCCCTTTCCCGCTTGCAAAATAGACATTTCTCAAAATCCAA 411
DB 205378 ATACCGCTGCAAGAGTCACTCTTCCCTGGTCTGTGAATAGACATTTCTCAAAATCCAA 205319
QY 412 AATGCCAGCAAGACCCCAATTTACTGAAACAGCCCAATACAGAAAGAAAGAAAT 471
DB 205318 AATGCTGCAAGACCCCAATTTACCTGAAAGCGGTACAAACAAGAAAGGAAAGAAAT 205259
QY 472 TAAACTGAGGACATTTCTCTCTGATATGATGATGATGATGATGATGATGATGATGATGAT 531
DB 205258 TAAACTGAGGACATTTCTCTCTGATATGATGATGATGATGATGATGATGATGATGATGAT 205199
QY 532 CACCATCCATTTGGCAAGAGGGCCAGCAGCATGCTTTTGGAGATATTTCTCTTTCTCA 591
DB 205198 CACTATACATCGGCAAGAGGGCCAGCAGCATGCTTTTGGAGATATTTCTCTTTCTCA 205139
QY 592 AGGGAAGTACGAGCTTTTACTGGAACAGGAGAAAGACACACTGGGGCAGTTCCTGG 651
DB 205138 AGGGAATTTATGAGCTTTGCCGGGAACCAAGAGAGGCACACTCTGGGCCAGTTCCCGG 205079
QY 652 GCATAATGATTTCTCCGGGCCCAACAGCACCTCGGACTCTGTGTTACAGAAACGCTCTC 711
DB 205078 GCATAACGACTTTCTCCGGGCCCAACAGCACCTCGGACTCAATGTTACAGAAACACCTC 205019
QY 712 CCGGGTGCCTCAAAATGCCATCTCCCTCCGACCAATGGAGGATCCCAAGCTCTCATGTT 771
DB 205018 CCGGGTGCCTCAAAATGCCATCTCCCTCCGACCAATGGAGGATCCCAAGCTCTCATG 204959
QY 772 GCCTTATTTGTCACGAGTACATTTAATTCGAACAGGAGTCTTCGGGCCAGCAAGCT 831
DB 204958 GCCTTATTTGTCACGAGTACATTTAATTCGAACAGGAGTCTTCGGGCCAGCAAGCT 204899
QY 832 GCCCAGGCTTAGCTGGAGCCGCTCATGGAGGAAAAGCTCAGGAGAAAAGCAGTCTGTT 891
DB 204898 GCCAAGGCTCAGCTGTGAGCCTGTATGGAAGAAAGTTTCAGGAGCAAAAGTAGTCTGTT 204839
QY 892 GGAAATGGGACAGTCCACAGGAGACACCTCGTGGGCTCCAGGGCTTCTGCATCTCA 951
DB 204838 AGAAGTGGGCAAGTCCACAGGAGACACCTCGTGGGCTCCAGTGGTTCGGGCTCTCA 204779
QY 952 GTCCAGCAAGGACAGACAGGACCTCTCCAGCCCTGTCGGAACAGTACCCGAGCTGGCC 1011
DB 204778 GTCCAGCAAGGACAGGACAGGACCTCTCCAGCCCTGTCGGAACAGTACCCGAGCTGGCC 204719
```

QY 2033 CTGATGATGCTTCTGCAGACAGTGGCTACACACATATGTTATTTGGAAATATAGT 2092
Db 203653 -----TGATGCTTCTGAGAGCTG-----GAGGTGAGTGTAGT 203620
QY 2093 ATTTTGTAGTACAGTACATCTTGCATACATAGTACTTCAACACACACATATAAAGTAA 2152
Db 203619 ATTTAGGACAGCTCAATGTTTCAATACATTTGGTTCATATCCACGCAATAAAGTAA 203560
QY 2153 ATGATAAGTGAACCTTGGTGTGTTATAGTAATAACAAGA--CCATAAGAGATAAGTAT 2210
Db 203559 ATGCTAAATGAACGGTTCATTTATAGTCATAACAGAACCGAAAGATTTTAAGACC 203500
QY 2211 AGCTAGAGAAATGCTTCTCTCTGAAATGTCATAGAGCCCTTAAGGTAAGAGATGATTTCCA 2270
Db 203499 ATCTAGTAAGTCTTAAATGTTATGACCCACCTTTAACAATAAGCAAGAGATGACTGCAG 203440
QY 2271 TCTACTCTCATTT 2283
Db 203439 TCTTGACTTTT 203427

AC055705 292420 bp DNA linear HTG 01-JUN-2002
Mus musculus clone RP23-17H14 strain C57BL6/J, *** SEQUENCING IN
PROGRESS ***, 91 unordered pieces.
AC055705
AC055705.11 GI:21306615
HTG: HTGS_PHASE1.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 292420)
Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
Gordon, M., Goltz, J.S. and Kucherlapati, R.
High Throughput Mouse Sequencing
Unpublished
2 (bases 1 to 292420)
Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
Gordon, M., Goltz, J.S. and Kucherlapati, R.
Direct Submission
Submitted (18-APR-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
3 (bases 1 to 292420)
Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
Gordon, M., Goltz, J.S. and Kucherlapati, R.
Direct Submission
Submitted (01-JUN-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
On Jun 1, 2002 this sequence version replaced gi:17976429.
-----Genome Center
Center: Harvard Partners Genome Center
Center Code: HPOC
Web site: <http://www.hpcgg.org/Sequence/mouse.html>
Contact: hpcg@emdel.mgh.harvard.edu
-----Summary Statistics
Center project name: ACF
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 100*
*Consensus quality: 274929 at least Q20
*Consensus quality: 267475 at least Q30
*Consensus quality: 233492 at least Q40
Estimated insert size: agarose-FP - N/A
**Estimated insert size: 290620 - sum-of-contigs
Quality coverage: agarose-FP - N/A
Quality coverage: 3.7 x in Q20 bases; sum-of-contigs estimation

NOTE: This is a 'working draft' sequence. It currently consists of 91 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 29525: contig of 29525 bp in length
29526: gap of unknown length
29546: contig of 19485 bp in length
49031: gap of unknown length
49031: contig of 15724 bp in length
64774: gap of unknown length
64775: contig of 11945 bp in length
76739: gap of unknown length
76740: contig of 11167 bp in length
76760: gap of unknown length
87927: contig of 11519 bp in length
87947: gap of unknown length
99466: contig of 11300 bp in length
110785: gap of unknown length
110805: contig of 11228 bp in length
122033: gap of unknown length
122034: contig of 11356 bp in length
133409: gap of unknown length
133410: contig of 6849 bp in length
140278: gap of unknown length
140298: contig of 8225 bp in length
148523: gap of unknown length
148543: contig of 7753 bp in length
148544: gap of unknown length
156297: contig of 6869 bp in length
163185: gap of unknown length
163186: contig of 6441 bp in length
163206: gap of unknown length
169667: contig of 4961 bp in length
174627: gap of unknown length
174628: contig of 5577 bp in length
180224: gap of unknown length
180244: contig of 5760 bp in length
186004: gap of unknown length
186025: contig of 4903 bp in length
190927: gap of unknown length
190947: contig of 4465 bp in length
195412: gap of unknown length
195413: contig of 4495 bp in length
199927: gap of unknown length
199947: contig of 4246 bp in length
204193: gap of unknown length
204213: contig of 4637 bp in length
208850: gap of unknown length
208870: contig of 1873 bp in length
210743: gap of unknown length
210764: contig of 2943 bp in length
213707: gap of unknown length
216378: contig of 2652 bp in length
216379: gap of unknown length
219944: contig of 3546 bp in length
219945: gap of unknown length
219965: contig of 1679 bp in length
221644: gap of unknown length
221663: contig of 1377 bp in length
223040: gap of unknown length
223041: contig of 1559 bp in length
223061: gap of unknown length
224639: contig of 1986 bp in length
226625: gap of unknown length
226626: contig of 85 bp in length
226730: gap of unknown length
226751: contig of 1380 bp in length
228131: gap of unknown length
228151: contig of 1294 bp in length

Tue Jul 29 09:34:14 2003

```
* 229445 229464: gap of unknown length
* 229465 230883: contig of 1419 bp in length
* 230884 230903: gap of unknown length
* 230904 232624: contig of 1721 bp in length
* 232625 232645: gap of unknown length
* 232646 233885: contig of 1241 bp in length
* 233886 233905: gap of unknown length
* 233906 235696: contig of 1791 bp in length
* 235697 235716: gap of unknown length
* 235717 236347: contig of 631 bp in length
* 236348 236367: gap of unknown length
* 236368 237362: contig of 995 bp in length
* 237363 237382: gap of unknown length
* 237383 238515: contig of 1133 bp in length
* 238516 238535: gap of unknown length
* 238536 239406: contig of 871 bp in length
* 239407 239426: gap of unknown length
* 239427 240632: contig of 1206 bp in length
* 240633 240652: gap of unknown length
* 240653 243078: contig of 2426 bp in length
* 243079 243098: gap of unknown length
* 243099 244569: contig of 1471 bp in length
* 244570 244589: gap of unknown length
* 244590 245470: contig of 881 bp in length
* 245471 245490: gap of unknown length
* 245491 247032: contig of 1542 bp in length
* 247033 247052: gap of unknown length
* 247053 248255: contig of 1203 bp in length
* 248256 248275: gap of unknown length
* 248276 249123: contig of 848 bp in length
* 249124 249143: gap of unknown length
* 249144 250545: contig of 1402 bp in length
* 250546 250565: gap of unknown length
* 250566 252352: contig of 1787 bp in length
* 252353 252372: gap of unknown length
* 252373 253486: contig of 1114 bp in length
* 253487 253506: gap of unknown length
* 253507 254903: contig of 1397 bp in length
* 254904 254923: gap of unknown length
* 254924 256191: contig of 1268 bp in length
* 256192 256211: gap of unknown length
* 256212 257056: contig of 845 bp in length
* 257057 257076: gap of unknown length
* 257077 258483: contig of 1407 bp in length
* 258484 258503: gap of unknown length
* 258504 259713: gap of unknown length
* 259694 259712: contig of 799 bp in length
* 259714 260513: gap of unknown length
* 260514 261603: contig of 1071 bp in length
* 260533 261623: gap of unknown length
* 261604 262443: contig of 820 bp in length
* 262444 262463: gap of unknown length
* 262464 263276: contig of 813 bp in length
* 263277 263296: gap of unknown length
* 263297 264251: contig of 955 bp in length
* 264252 264271: gap of unknown length
* 264272 264730: contig of 459 bp in length
* 264731 264750: gap of unknown length
* 264751 266098: contig of 1348 bp in length
* 266099 266118: gap of unknown length
* 266119 267211: contig of 1093 bp in length
* 267212 267231: gap of unknown length
* 267232 268372: contig of 1141 bp in length
* 268373 268392: gap of unknown length
* 268393 269932: contig of 1540 bp in length
* 269933 269952: gap of unknown length
* 269953 271829: contig of 1877 bp in length
* 271830 271849: gap of unknown length
* 271850 272547: contig of 698 bp in length
* 272548 272567: gap of unknown length
* 272568 273748: contig of 1181 bp in length
* 273749 273768: gap of unknown length

* 273769 274659: contig of 891 bp in length
* 274660 274679: gap of unknown length
* 274680 275615: contig of 936 bp in length
* 275616 275635: gap of unknown length

Query Match 35.2%; Score 895.6; DB 2; Length 292420;
Best Local Similarity 68.9%; Pred. No. 4e-207;
Matches 1498; Conservative 0; Mismatches 559; Indels 117; Gaps 15;

QY 175 CCGTTTCTCTGGTTGGTCAAGACTGGATGATGTAACCTGGCTCTCTAGGAAGCTCA 234
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74613 CAGGTTTCTCTGAGCTGTGAAGGACTCGATAATGTAACCTGGCTCTCCGGAAGACCCG 74672
QY 235 CTTGGCCGTAACTCAGGAAGTCTCTTTTGCACCCCATCTCATTTTGAAGACCACTTCTGA 294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74673 CTTGGCCCTTGGCTCAGCAAGATCTCAAGACCCCTGCTCTGTTTTAGAGCCACTTGAGA 74732
QY 295 AGCCACTTTGAGAAAATGATGTGACAGTTCCTATCAAAAAGGA---TTCAGAAACATATA 351
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74733 AG-----AAATGATGTGACATTTCTATGAAPAGGAGCTTTTGGAAAGTGAAG 74781
QY 352 CCATCTGTGAAGAAAGTGGCCCTTTCTCCCGCTTGCAAAATAGACATTTCTCAAAATCCAA 411
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74782 ATACCGGTGCAAGAGGTCACTCTTTCCCTGGTGTGTAATAGACATTTCTCAAAATCCAA 74841
QY 412 AATGCCAGCAAGACCCCAATTTACCTGAAAGACCCCAATTAACAGAAAGAAAGAAATTT 471
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74842 AATGCCCTGCCAAGACCCCAATTTACCTGAAAGCGCTAACACAAAGAAAGAAAGAAATTT 74901
QY 472 TAAACTGAGGGACATTTCTCTCTCTGATATGATAGTCCCGCTTGGAGACTTTTCGGCA 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74902 TAAACTGAGGGACATTTCTCTCTCTGATATGATAGTCCCGCTTGGAGACTTTTCGGCA 74961
QY 532 CACCATCCACATTTGGCAAGAGGGCCAGCAGCATGTCTTTTGGAGATATTTCTTCTTTCA 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74962 CACTATACACATCGGCAAGAGGGCCAGCATGACGCTTTTGGGGATATTTCTTCTTTCA 75021
QY 592 AGGAACTAGAGCTTTTACCTGGAACACAGAGAAAGACACCTTGGCCAGTTCCTCTGG 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75022 AGGAACTATGAGCTTTTCCCGGGAACCAAGAGAGGACACACTTGGCCAGTTCCTCTGG 75081
QY 652 GCATATAGTCTTCTCCGGGCCAACAGCAGCTCGGACTCTGTGTTCACAGAAACCCCTC 711
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75082 GCATAGGACTTCTTCGGGCCAACAGCAGCTCGGACTCAATGTTTACAGAAACCCCTC 75141
QY 712 CCGGTGCTCAAAATGCCATCTCCCTCCGACCATTTGGAGGATCCCAAGCTCTCATGTT 771
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75142 CCGGTGCTCAAAATGCCATCTCCCTCCGACCATTTGGGGATCCCAAGCTCTCATGCT 75201
QY 772 GCCCTTATGTCACCACTGACATTTTAAATCCAAACAGGAGTCTTCCGGCCAGCAAGCT 831
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75202 GCCCTTATGTCACCACTGACATTTTAAATCCAAACAGGAGTCTTCCGGCCAGCTTCAAGTT 75261
QY 832 GCCCAGGCTTAGCTGCGAGCCGCTCATGAGGAGAAAGCTCAGGAGAAAGCAGTCTGTT 891
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75262 GCCAAGCTCAGCTGTGAGCTGTGATGGAAGAGAAAGTTTCAAGGACAAAGTAGTCTGTT 75321
QY 892 GGAGAAATGGGACAGTCCACAGGGAGACACTCGTGGGGCTCCAGCGTCTCTCATCTCA 951
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75322 AGAGAAATGGGACAGTCCACAGGGAGACACTCGTGGGGCTCCAGTGGTTTCCGGGTCTCA 75381
QY 952 GTCCAGCCAAAGCAGAGACAGCCACTCTCTCAGGCTTCCGAAACAGTACCCGACCTGGCC 1011
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75382 GTCCAGCCAAAGCAGCCGAGACAGCCACTCTCTCAGGCTTCCGAAACAGTACCCGACCTGGCC 75441
QY 1012 AGCCGAGACATGTTTGGACCATCCACCCCATCGGAGCTCATCAAGAGAAAGACTAAGTC 1071
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75442 CGCGACGACATGTTTGGAGCATCTCTGCTCTGTGAGCTTGTGAAGTCAAGAACTAAATC 75501
QY 1072 AGAGAGTCCCTCTCTGACCTTACAGGTTCCCTCTCTCCTCGAGCTTGTATCTTGGGCC 1131
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75502 AGAAGAGTCTTCTCTGACCTGACCGGCTCCCTCTCTCTCCTCGAGCTTGTATCTTGGGCC 75561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


Query Match 30.2%; Score 769.8; DB 10; Length 1481;
Best Local Similarity 75.2%; Pred. No. 7.6e-177;
Matches 1062; Conservative 0; Mismatches 297; Indels 54; Gaps 6;

QY 20 CGGCCACCGCGCTGGCGCTTCACGCCCTCGGAGCGGATAATGCGTGGAGCAGCA 79
DB 60 CGGCTGCTCCCGGAGACCGCGCGAGCTCCGCGAGCAGATCTCGCGGTGAACCGTTG 119
QY 80 CCAGCGCGGAGACTCGGCTGGATCTCG-----CACAGCGGAGGAGTTCGGT 128
DB 120 CGGCCCTCTGTTCTCGCGCACCGAGGAGGACTGTTCGCCCTCGGAGGCGACCCCG 179
QY 129 GCGCCCGGAGGCGCGCGCGAGCGGCTGGGATCTCTAGCGCGCGCGGTTTGTCTCG 188
DB 180 GCGGACCGCGCGCTCCCGGAGAGAGCTAGAGAGCTGCTGTTCTCCGAG 239
QY 189 TTGGGTCAGACTGGATGATGTAAGTGGCTCTTAGGAAGCTCTACTTGGCGGTAACT 248
DB 240 CTGTGGAAGGAGACTCGATAATGTAAGTGGCTCTCCGGAAGAGACCGCGTGGCGCT 299
QY 249 CAGGAAGTCTCTTTGACCCCATCTCATTTCAAGCCACTCTGAAGCCACTTGAGAA 308
DB 300 GAGCAAGATCTCCAGACCTGTCTCTGTTTATAGGCCACTTGAGAG-----A 348
QY 309 AATGATGTGACAGTTCCTATCAAAAAGGA---TTCAGAAACATATACCATCTGTGAAG 365
DB 349 AATGATGTGACATCTCTATGAAAAGGAGCTTTTGAAGTGTAGGATACCCGTGCAG 408
QY 366 AGTGGCGCTTTCGCGCTTTCCTATCAAAAAGGA---TTCAGAAACATATACCATCTGTGAAG 425
DB 409 GTTCTGCTCCCTGATATGATCAGTCTCCCTCGCGGAGCTTTCGTCACATATACATCG 588
QY 426 CCCCAATTTACCTGAAGAGCAATCAAGAAAGAAAGAAATTAACCTGAGGACA 485
DB 469 CCCCAATTTACCTGAAGAGCAATCAAGAAAGAAAGAAATTAACCTGAGGACA 528
QY 486 TTTCTGCTCTGATGATCAGTCCCGCTTGGAGACTTTCGCCACCATCCACATG 545
DB 529 TTTCTGCTCCCTGATATGATCAGTCTCCCTCGCGGAGCTTTCGTCACATATACATCG 588
QY 546 GCAAGAGGCGCAGCAGGATGCTTTGGAGATATTCCTTCTTCAAGGAACTAGCAGC 605
DB 589 GCAAGAGGCGCAGCAGGATGCTTTGGAGATATTCCTTCTTCAAGGAAATATGAGC 648
QY 606 TTTTACCTGGAACACAGGAGAAAGCACACCTGGCGGAGTTCCTTGGGAGTATGAGTCT 665
DB 649 TCTTGGCGGAAACAGAGAGAGGACACTCTGGCCAGTTCGCCGGGAGTATGAGTCT 708
QY 666 TCGGGGCAACAGCAGCTCGGACTCTGTGTTACAGAAAGCGCCCTCCCGGCTCAAAA 725
DB 709 TCGGGGCAACAGCAGCTCGGACTCTGTGTTACAGAAAGCGCCCTCCCGGCTCAAAA 768
QY 726 ATGCCATCTCCCTCCGACCATTTGGAGGATCCCAAGCTCTCATGTTGCCCTTATGTCAC 785
DB 769 ATGCCATCTCCCTCCGACCATTTGGAGGATCCCAAGCTCTCATGTTGCCCTTATGTCAC 828
QY 786 CAGTGACATTTAATTCCAAACAGGAGTCTCTTGGCGGAGCAAGAGTCCGAGGCTTAGCT 845
DB 829 CAGTGACATTTAATTCCAAACAGGAGTCTCTTGGCGGAGTCTTGAAGTTCGCAAGCTCAGCT 888
QY 846 GCGAGCGCGTCAATGGAGGAAAGAGCTCAGGAGAAAGAGCTCTGTTGGAGAAATGGGACAG 905
DB 889 GTGAGCGCTGTCATGGAAGAGAAAGTTCAGGAGCAAAAGTCTATTAGAGAAATGGGCGAG 948
QY 906 TCCACAGGAGAGACACTGTGGGGTCCAGCGGTTCTGCATCTCAGTCCAGCCAGGCA 965
DB 949 TCCACAGGAGAGACACTGTGGGGTCCAGCTGTTCCGGGCTCTAGTCCAGCCAGGCGC 1008
QY 966 GAGACAGCACTCTCCAGCTCTCGGAACAGTACCCGACCTGCGGAGCGGAGGAGCATCT 1025
DB 1009 GCGACAGCACTCTCCAGCTCTCGGAACAGTCTCCGAGCTGGCGCGGAGGAGCATCT 1068

QY 1026 TTGACATCCACCCCATCGAGCTCATCAAGGAGAAAGACTAAGTCAGAGAGTCCCTCT 1085
DB 1069 TTGAGATCTCCCTCCCTCTGAGCTTGGAAGTCAAGAGCTAAATCAGAAGAGTCTCTCT 1128
QY 1086 CTGACCTTACAGGTTCCCTCTCTCCCTGACGCTTGTGATCTTGGGCGCTCAGCTTTGGATG 1145
DB 1129 CTGACCTGACCGGCTCCCTGCTCTCCCTGACGCTTGTGATCTTGGGCGCTCAGCTTTGGATG 1188
QY 1146 AGTGTCTGAATGATGATGATAAATAAGTAACAGATGCCAAGCTTTTCTTTGGGCT 1205
DB 1189 AAGTCTGATGATGATGATAAATAAGTAGCCAGAGTCAAGCTCTTTTCTTTGGGAGT 1248
QY 1206 AAAAGGTAC-----AAAAAACAATAACACAGTTGAAGAGAGAGG 1246
DB 1249 AGAGGTACCAATAAAGCAGAAACAAAAAATCACTAACACAGCTGATGGAAGAG 1308
QY 1247 CTTCCGGAGCTG-----TATTTGACGTTTGTGTGGGTTTCTTAAATAA-----TATT 1296
DB 1309 CTGCTCTGACTGTTTATTTATTTGACGCACTGTGATCATTTTCTTAAATAAAGCTCTCTGT 1368
QY 1297 CTTCAAGATATTTTACCTGTTATGCCCTGTTTGCAGAAACAAATTTAGAAAAACA 1356
DB 1369 CTTCAAGATATTTTATTAACCATCATGCCCTGTTTGCAGAAATTTTGAACAATAAT 1428
QY 1357 ACAAGCAAAACCTATCTTGGCAAAAAAGGAA 1389
DB 1429 TTTAAAAAATAAATAAATAAATAAATAAATAA 1461

RESULT 10
AF094521
LOCUS Homo sapiens MSE55-related protein (UBI) mRNA, complete cds.
DEFINITION AF094521
ACCESSION AF094521.1 GI:3834632
VERSION
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 765)
AUTHORS Alberts, A.S., Bouquin, N., Johnston, L.H. and Treisman, R.
TITLE Analysis of RhoA-binding proteins reveals an interaction domain conserved in heterotrimeric G protein beta subunits and the yeast response regulator protein Skn7
JOURNAL J. Biol. Chem. 273 (15), 8616-8622 (1998)
MEDLINE 98204843
PUBMED 9535835
REFERENCE 2 (bases 1 to 765)
AUTHORS Alberts, A.S.
TITLE MSE55-related Cdc42-binding protein
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 765)
AUTHORS Alberts, A.S.
TITLE Submitted
JOURNAL Submitted (24-SEP-1998) Cancer Center, University of California, San Francisco, 2340 Sutter St., San Francisco, CA 94115, USA

FEATURES
source
1..765
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="brain"
/dev_stage="fetal"
1..765
/gene="UBI"
1..765
/note="Cdc42 binding protein; similar to Homo sapiens serum protein MSE55: SwissProt Accession Number Q00587 and unknown protein encoded by the sequence presented in Genbank Accession Number AF001436; contains CRIB-motif"
/codon_start=1
/product="MSE55-related protein"

```
/protein_id="AAC71773.1"
/db_xref="GI:3834633"
/translation="MPAKTPIYLKAANNKKKFKLRLDILSPDMISPLGLDFRHTIHI
KGEQHDVFGDISFLOGNYELLPNGOKAHLQFPNGHFRANSTSDSVETPSPV
LKNALISLPTIGDSQALMLPLSPVTFNSKQESFGPAKLPLRSLCEPVMEKRAQKSSLL
ENGIVHQGDTWSSGSSASQSSQGRSHSSLSQYPDWPAEDMFDPHTPCELIKGT
KSEESLDTGLSLFLQLDLGSLDLVENVMDKNK"
BASE COUNT      203 a      221 c      178 g      163 t
ORIGIN

Query Match      29.9%; Score 761.8; DB 9; Length 765;
Best Local Similarity 99.7%; Pred. No. 6.1e-175;
Matches 763; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 413 ATGCAGGCAAGACCCCAATTTACCTGAAGCAGCAATCAACAAGAAAGAAATTT 472
    |||||
DB 1 ATGCCAGCAGACCCCAATTTACCTGAAGCAGCAATCAACAAGAAAGAAATTT 60
    |||||

QY 473 AAACGTGAGGACATTTCTCTCTGATATGATCATGTCCTCCCGCTTGAGACATTTTCGCCAC 532
    |||||
DB 61 AAACGTGAGGACATTTCTCTCTGATATGATCATGTCCTCCCGCTTGAGACATTTTCGCCAC 120
    |||||

QY 533 ACCATCCACATGGCAAGAGGGCCAGCAGCATGCTTTGGAGATATTTCTTTCTTCAA 592
    |||||
DB 121 ACCATCCACATGGCAAGAGGGCCAGCAGCATGCTTTGGAGATATTTCTTTCTTCAA 180
    |||||

QY 593 GGGAACTACGAGCTTTTACCTGGAAACCCAGGAGAAACACACCTGGGCCAGTTCCTCTGGG 652
    |||||
DB 181 GGGAACTACGAGCTTTTACCTGGAAACCCAGGAGAAACACACCTGGGCCAGTTCCTCTGGG 240
    |||||

QY 653 CATAATGATGTTCTCCGGGCCCAACAGCACCTCGGACTCTGTGTTTCACAGAAACGCCCTCC 712
    |||||
DB 241 CATAATGATGTTCTCCGGGCCCAACAGCACCTCGGACTCTGTGTTTCACAGAAACGCCCTCC 300
    |||||

QY 713 CCGGTGCTCAAAATGCCATCTCCCTCCGACCATTGAGCATCCAGACTCTCATGTTG 772
    |||||
DB 301 CCGGTGCTCAAAATGCCATCTCCCTCCGACCATTGAGCATCCAGACTCTCATGTTG 360
    |||||

QY 773 CCCTTATTTGTCACAGTACATTTAAATTCAAACAGGAGTCTCTCGGGCCACGAAAGCTG 832
    |||||
DB 361 CCCTTATTTGTCACAGTACATTTAAATTCAAACAGGAGTCTCTCGGGCCACGAAAGCTG 420
    |||||

QY 833 CCCAGGCTTAGTGGAGCCCGCTCATGGAGGAAAGCTCAGGAGAAAGCAGTCTGTTG 892
    |||||
DB 421 CCCAGGCTTAGTGGAGCCCGCTCATGGAGGAAAGCTCAGGAGAAAGCAGTCTGTTG 480
    |||||

QY 893 GAGATGGSACAGTCCACCAGGAGACACCTCGTGGGCTCCAGCGTCTGCAATCTCAG 952
    |||||
DB 481 GAGATGGSACAGTCCACCAGGAGACACCTCGTGGGCTCCAGCGTCTGCAATCTCAG 540
    |||||

QY 953 TCCAGCCAGGAGAGAGAGAGGAGTCTCTCCAGGCTGTCCGAAACAGTACCCGAGTGGCCA 1012
    |||||
DB 541 TCCAGCCAGGAGAGAGAGAGGAGTCTCTCCAGGCTGTCCGAAACAGTACCCGAGTGGCCA 600
    |||||

QY 1013 GCCGAGGACATGTTGACATCCACCCATCCGAGTCTATCAGGAGAAAGCAATGTC 1072
    |||||
DB 601 GCCGAGGACATGTTGACATCCACCCATCCGAGTCTATCAGGAGAAAGCAATGTC 660
    |||||

QY 1073 GAGGAGTCCCTCTGACCTACAGGTTCCCTCTCTCCCTGAGCTTGATCTTGGGCC 1132
    |||||
DB 661 GAGGAGTCCCTCTGACCTACAGGTTCCCTCTCTCCCTGAGCTTGATCTTGGGCC 720
    |||||

QY 1133 TCACCTTTTGGATGAGTCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1177
    |||||
DB 721 TCACCTTTTGGATGAGTCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 765
    |||||

RESULT 11
AC090491/c      204019 bp      DNA      linear      HTG 24-FEB-2001
LOCUS
DEFINITION      Mus musculus clone RP23-12K1, *** SEQUENCING IN PROGRESS ***, 43
unordered pieces.
AC090491
```

VERSION
KEYWORDS
SOURCE
ORGANISMREFERENCE
AUTHORSTITLE
JOURNAL
AUTHORS
TITLE
JOURNAL

COMMENT

AC090491.1 GI:13122740

HTG; HTGS_PHASE1.

Mus musculus.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 204019)

McCombie, W.R., Baker, J.P., Bahret, A., Bal, H., Balija, V.,

Dedhia, N.N., de la Bastide, M., Huang, E.N., King, L., Kirchoff, K.A.,

Miller, B., Nascimento, L.U., O'Shaughnessy, A.L., Preston, R.R.,

Rodriguez, S., Santos, L., Shah, R.S., Spiegel, L.A., Toth, K., Vil, M.D.

and Zutavern, T.

Mouse Genomic Sequence

Unpublished

2 (bases 1 to 204019)

McCombie, W.R.

Direct Submission

Submitted (24-FEB-2001) Lita Annenberg Hazen Genome Sequencing

Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring

Harbor, NY 11724, USA

----- Genome Center

Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor

Laboratory

Center code: CSHL

Web site: <http://www.cshl.org/genseq>Contact: mccombie@cshl.org

----- Project Information

Center project name: RP23-12K1

Center clone name: RP23-12K1

* NOTE: This is a 'working draft' sequence. It currently

* consists of 43 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 12288: contig of 12288 bp in length

* 12289 12369: gap of unknown length

* 12370 24613: contig of 12244 bp in length

* 24614 24694: gap of unknown length

* 24695 36517: contig of 11823 bp in length

* 36518 36599: gap of unknown length

* 36599 45719: contig of 9121 bp in length

* 45720 45800: gap of unknown length

* 45801 54340: contig of 8540 bp in length

* 54341 54421: gap of unknown length

* 54422 62884: contig of 8463 bp in length

* 62885 62965: gap of unknown length

* 62966 70952: contig of 7987 bp in length

* 70953 71033: gap of unknown length

* 71034 77688: contig of 6655 bp in length

* 77689 77770: gap of unknown length

* 77770 84397: contig of 6628 bp in length

* 84398 84478: gap of unknown length

* 84479 90811: contig of 6333 bp in length

* 90812 90892: gap of unknown length

* 90893 96814: contig of 5922 bp in length

* 96815 96896: gap of unknown length

* 96896 102124: contig of 5229 bp in length

* 102125 102205: gap of unknown length

* 102206 107275: contig of 5070 bp in length

* 107276 107356: gap of unknown length

* 107357 112299: contig of 4943 bp in length

* 112300 112380: gap of unknown length

* 112381 117254: contig of 4874 bp in length

* 117255 117336: gap of unknown length

* 117336 122208: gap of unknown length

* 122208 126863: contig of 4575 bp in length

* 126864 126944: gap of unknown length

* 126945 131059: contig of 4125 bp in length

LOCUS	BC022337	1670 bp	linear	PRI 04-FEB-2002
DEFINITION	Homo sapiens, Cdc42 effector protein 2, clone MGC:23686			
ACCESSION	IMAGE:4690444, mRNA, complete cds.			
VERSION	BC022337			
KEYWORDS	BC022337.1 GI:18490847			
SOURCE	MGC.			
ORGANISM	Homo sapiens.			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1670)			

QY 763 TCTCATGTGCCC 775
 |||||
Db 441 TCTCACCCTGCCC 453

Search completed: July 25, 2003, 18:27:13
Job time : 6716.77 secs

